





XX A method has been developed for the diagnosis of human rheumatoid arthritis (RA) using two-dimensional electrophoresis to generate a two-dimensional array of features. The method can be used for screening, diagnosis and prognosis of RA in a subject or for monitoring the effect of an anti-RA drug or therapy administered to a subject. The method comprises: (a) analysing a sample of serum or plasma and optionally synovial fluid by two-dimensional electrophoresis, to generate a two-dimensional array of features; (b) identifying with at least one chosen feature whose relative abundance correlates with the presence or absence of RA; and (c) comparing the abundance of each chosen feature in the sample with the abundance of that chosen feature in serum or plasma from one or more persons without RA, where the relative abundance of the chosen feature or features in the sample indicates the presence or absence of RA in the subject. The method can also be used in clinical studies for testing drugs for therapy of RA, for purification of RA diagnostic protein isoforms (RPIs), and for production of antibodies to RPIs. The RA-diagnostic feature (DAF) proteins can be used to identify compounds that promote or inhibit their activity, which are then used as RA drugs. Nucleic acid encoding RPIs can be used in gene therapy protocols. AA44184 to AA44200 represent RPI peptides, AA44201 to AA44206 represent expression vectors, protein isoform peptides and AA44206 to AA44208 represent peptide-antibody pairs, which are all used in the exemplification of the present invention.

Sequence: 16 AA:

Query Match 100.0%; Score 72; DB 20; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5.4e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	PPHPLSHGSAQV	13
Pb	3	PPHPLSHGSAQV	15

## RESULT 4

TD	AAW66275 standard; peptide; 19 AA.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AC: AAW66275;

DT 23-NOV-1998 (first entry)

DE: Circularly permuted di-alpha haemaglobin-like peptide fragment.

KW haemagglutinin; circularly permuted; N-terminus; C-terminus; multimeric;

$$s_1 + \dots + s_k \mid a \iff s_1 + \dots + s_k \mid a.$$

Synthetic

PN W09838211-A2

PD 03-SEF-1998.

PI 27-FRH-1948; 98WO-US03836.

PR 28-FEB-1997; 97HS-0039657.

PA (SOMA-); SOMATOGEN INC.

PI Best PA, Collins P.;  
YY

DR WP1; 1998-481141/41.  
XY

PT New haemostatic fermiesin with non-naturally occurring termini -  
 PT useful for the treatment of any condition where therapy includes red  
 blood cells or oxygen delivery  
 XX  
 Disclousure Date: 10/27/91; Eclairsh.

The invention relates to a haemoglobin permutoxin (pHb), containing at least one substitution with non-ferratingly occurring N and C terminal. Also claimed is a nucleic acid encoding the above pHb. pHb can be used as a substitute for red blood cells whenever red blood cells or oxygen delivery is required, particularly in the treatment of haemorrhages, traumas, and surgeries requiring blood transfusion or reinfusion of the patient's own blood and anaemia, to deliver oxygen to tissue sites that red blood cells cannot penetrate, including site downstream of an obstruction as in thrombi, stroke cell occlusions, arterial occlusions, angioplasty balloons, and surgical instrumentations, and to treat tissue ischaemias including stroke, emerging stroke, transient ischaemic attacks, myocardial stunning and hibernation, acute or unstable angina, and to treat infarct. The invention can also be used to deliver drugs and for *in vivo* imaging. The permutoxins have increased stability and expression during E. coli synthesis and a greater *in vivo* circulating half life compared to the listed hb molecules of prior art. The present sequence is shown in the specification.

SQ sequence: 19 AA;

Query Match	100.0%	Score 72	PR 14	Length 19
Best Local Similarity	100.0%	Prod. No. 0.0e+07		
Matches	13	Conservative	0	Indels
			0	Gaps
			0	

```
QY      1 FPHFDLSHGSAQY 13
          |||||
Db      6 fphfdlshtgag 18
```

## RESULTS

ID AAR29606 standard; Protein; 141 AA.

AC AAF29606;

DT 15-JAN-1993 (first entry)

DH: Mutant human alpha-globin 1.

KW Tetrahydropyran intramolecular crosslink; oxygen exchange; erythrocytes; HIV.

05 Synthetic

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

/\*taq- a

[illegible]XX  
00 MAY 1969XX  
00 1900 1601 0100 100066704[illegible]XX  
XX  
CIVIL - A UNITV AL AGAMA BEC EQUINO

XX  
XX  
PT  
McQuinn, El  
Toussou, TM

XX  
WB: 1660-250026-730

DR N-PSDB; AAQ26424.  
XX

transgenic, crosslinked polymeric human interleukin-1 (IL-1) as a cell-free blood substitute

PS Disclosure: Page 13: 4200: English

CC This sequence was not disclosed in the specification. The alpha  
CC platin gene sequence was obtained from Swiss-pat. no. AAP01922. The  
CC mutation was then added to the sequence. The sequence was mutated by  
CC polymerase chain reaction. The mutation introduced produced a

cc Polymer and molecular crosslink. This alpha molecule has oxygen  
 cc exchange characteristics attributed to its use to natural hemoglobin (Hb)  
 cc to enable its use in a cell free blood substitute. The mutated  
 cc gene encoding this protein can be used in a construct designed for  
 cc expression of pure modified human Hb's in the cythology of non  
 cc human transgenic animals. The Hb's are non immunogenic so can be used  
 cc without the need for blood typing. Also, the transgenic Hb will be  
 cc free of such viruses as HIV.  
 cc  
 cc Sequence: 141 AA

Query Match: 100.0%; Score: 72; DB: 14; Length: 141;  
 Best Local Similarity: 100.0%; Prod. No. 6, 6, 06;  
 Matches: 14; Conserved: 0; Mismatches: 0; Indels: 0; Gaps: 0;

07 1 FPHHSHSAAV 14  
 10 1111111111  
 1b 48 Ippdshsaaav 50

# RESULT 6

AAR29000 standard; Protein: 141 AA.

AAR29000

AAR29000 (First entry)

Medium human alpha globin 2

Human alpha globin 2

Human alpha globin 2

Key: Location/Qualifiers

PI MISC difference 74  
 11 22-DBP-1992

12 06-06T-1989; 800S-0417949;

13 06-06T-1989; 800S-0417949;

14 06-06T-1989; 800S-0417949;

15 06-06T-1989; 800S-0417949;

16 06-06T-1989; 800S-0417949;

17 06-06T-1989; 800S-0417949;

18 06-06T-1989; 800S-0417949;

19 06-06T-1989; 800S-0417949;

20 06-06T-1989; 800S-0417949;

21 06-06T-1989; 800S-0417949;

22 06-06T-1989; 800S-0417949;

23 06-06T-1989; 800S-0417949;

24 06-06T-1989; 800S-0417949;

25 06-06T-1989; 800S-0417949;

26 06-06T-1989; 800S-0417949;

27 06-06T-1989; 800S-0417949;

28 06-06T-1989; 800S-0417949;

29 06-06T-1989; 800S-0417949;

30 06-06T-1989; 800S-0417949;

cc Polymer and molecular crosslink. This alpha molecule has oxygen  
 cc exchange characteristics attributed to its use to natural hemoglobin (Hb)  
 cc to enable its use in a cell free blood substitute. The mutated  
 cc gene encoding this protein can be used in a construct designed for  
 cc expression of pure modified human Hb's in the cythology of non  
 cc human transgenic animals. The Hb's are non immunogenic so can be used  
 cc without the need for blood typing. Also, the transgenic Hb will be  
 cc free of such viruses as HIV.  
 cc  
 cc Sequence: 141 AA

Query Match: 100.0%; Score: 72; DB: 14; Length: 141;  
 Best Local Similarity: 100.0%; Prod. No. 6, 6, 06;  
 Matches: 14; Conserved: 0; Mismatches: 0; Indels: 0; Gaps: 0;

07 1 FPHHSHSAAV 14  
 10 1111111111  
 1b 48 Ippdshsaaav 50

# RESULT 7

AAR40543 standard; Protein: 141 AA.

AAR40543

AAR40543 (First entry)

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Human alpha globin

Db 43 fpndishsqap 55

RESULT 8  
ID AAR32988 standard: protein: 141 AA

AC AAR32988;  
DT 17-JUN-1993 (first entry)

DE Human alpha haemoglobin and petunia E5P synthase transit peptide.

XX Recombinant products; commercial production; fermentation;  
XX biosynthesis; natural products; recombinant proteins;  
XX product expression; protein expression; expressed proteins;  
OS Chimeric - Homo sapiens.  
OS Chimeric - Petunia sp.

XX WC9303161-A.

XX 18-FEB-1993.

XX 31-JUL-1992: 93WO-0596359.

XX 01-AUG-1991: 91US-0739143.

XX (DAMS/) JAWSON W Q.

XX (DONS/) GARGER S J.

XX (GRAN/) GRANTHAM G L.

XX (GRIL/) GRILLE L K.

XX (TUPP/) TUPPEN A M.

XX (TUPP/) TUPPEN T H.

XX Dawson W, Jenson T, Garger S, Grantham G, Grille L;

XX Turpen A, Turpen TH;

XX WPI: 1993-076518/09.

XX N-PSDB: Q376810.

XX Recombinant plant viral nucleic acids - used to express a prod.,

XX e.g. antibody or IL-1 in a plant

XX Example 4: Page 105; 40pp; English.

XX This sequence is human alpha-haemoglobin with petunia E5P synthase

XX transit peptide. The coding sequence was inserted into recombinant plant

XX viral nucleic acid which was then used to express a recombinant product

XX (in this case human alpha-Hb) in a plant. The plant viral sequence

XX may be from tobacco mosaic, cucumber green mottle, cowpea mosaic, bromo

XX mosaic, broad bean mottle, rice necrosis, geminiviruses, tomato golden

XX mosaic, cassava latent and maize streak viruses.

XX Sequence 141 AA.

XX Query Match 100.0%; Score 72; DB 14; Length 141;

XX Best Local Similarity 100.0%; Prod No 6, 3e-06;

XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Oy 1 PPHPSHSNACV 13

XX Db 43 fpndishsqap 55

XX RESULT 9

XX AAR42631

XX ID AAR42631 standard: protein: 141 AA.

XX AC AAR42631;

XX 26-APR-1994 (first entry)

XX Natural alpha-globin.

XX Haemoglobin; alpha globin; beta globin; blood substitute;

XX erythrocyte; erythrocyte; erythrocyte; erythrocyte; erythrocyte;

XX fusion protein.

XX WC9319089-A.

XX 30-SEP-1993.

XX 18-MAR-1993: 93WO-FR00273.

XX 18-MAR-1993: 93FE-0007224.

XX (INRM) INSEPM INST NAT SANTE & RECH MEDICALE.

XX (INRM) FASIEUK MERLUX SERUMS & VACCINS.

XX Edelstein S, Pascher RJ, Foyart G;

XX WPI: 1993-07682740

XX N-PSDB: AAG49615.

XX M & synthetic tetra- and globin type oxygen transporter - with

XX all chains identical and contg. haem, useful as blood substitute

XX for transfusion

XX Disclosure: Fig 1; 41pp; French.

XX Chimeric alpha beta globin molecules are claimed which, when

XX associated together to reconstitute the alpha-beta interface and

XX incorporating haem, are suitable as blood substitutes. Pref. the

XX chimeric chains contain amino acids 1-73 of beta globin at the

XX C terminus and amino acids 69 onwards of alpha-globin at the

XX C terminus. Alternatively, each chain may contain the N-terminus of

XX an alpha chain with the C-terminus of a beta-chain, or all 4

XX chains can be of the beta type. The tetramers are characterised by

XX a lower oxygen affinity than natural haemoglobin.

XX Sequence 141 AA;

XX Query Match 100.0%; Score 72; DB 14; Length 141;

XX Best Local Similarity 100.0%; Prod. No. 6, 3e-06;

XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Oy 1 PPHPSHSNACV 13

XX Db 43 fpndishsqap 55

XX RESULT 10

XX AAR55131

XX ID AAR55131 standard: protein: 141 AA.

XX AC AAR55131;

XX 12 JAN 1995 (first entry)

XX Human alpha haemoglobin.

XX Virus: recombinant plant virus; alpha triphosphatidyl phenol type;

XX alpha amylinase; alpha haemoglobin; bromo mosaic virus; gemini virus;

XX rice necrosis virus; tobacco virus; acute expression; Chinese cucumber;

XX Homo sapiens.

XX US5316941-A.

XX 31-MAY-1994.

```

11 26 FEB 1994 0005 01607566
XX
12 26 FEB 1994 0005 01607566
13 26 FEB 1994 0005 01607566
14 15 JUN 1994 0005 02192796
15 17 FEB 1994 0005 03109811
16 05 MAY 1994 0005 03476337
17 08 JUN 1994 0005 03643498
18 22 OCT 1994 0005 03600244
19 15 JUN 1994 0005 03419117
20 15 JUN 1994 0005 03478996
21 01 AUG 1994 0005 07091111
XX
22 (BLAST) BLASTOFF: GENETICS EXPLO.
XX
23 Lowman WB, Jones of L, Garner SD, Strehlan GL, CHILLER
24 Forster AM, Forster JM,
25 WELT 1994 17629 92/93
26 N PDBID: AAK44517
XX
27 New recombinant plant viral nucleic acid capable of systemic
28 infection and stable expression of non native nucleic acid in
29 plant host
30 Example 4: sequence 5' 383 4449: English
31
32 The alpha hemoglobin coding sequence was fused to a sequence
33 encoding the protein HSP synthase (heat peptide and inserted into
34 a recombinant plant virus which can then be used to infect plants
35 for the production of non native products (in this case human
36 alpha hemoglobin). Other genes which may be inserted into the
37 virus are those which control a phenotypic trait, such as anti
38 sterility or sequences encoding a sense RNA which can be used
39 to prevent the expression of undesired phenotypic traits. The
40 recombinant virus is derived from a plus sense, stable stranded
41 virus selected from rubomoviruses, bromovirus, tobacco etch virus
42 or virus of a similar virus.
43
44 Sequence 141 AA:
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
25
```

2R 12-JUN-1992: 92US-0897648.  
 2R 08-DEC-1992: 92US-0987890.  
 2K 15-MAR-1994: 94US-0030897.  
 XX  
 XX (DNXD-) DNK CORP.  
 XX  
 P1 Holzman SH, Kellor H, Kumar P, Logan JS, O'Connell JK;  
 P1 Parsons CT, Pilder SH, Pinkert CA, Sharma A, Swanson ME;  
 P1 White SP;  
 P1 WPI: 1994-007675/01.  
 XX  
 XX Transgenic pigs which produce human haemoglobin: conty. DNA encoding  
 P1 human alpha globin and huma beta globin linked to promoters  
 PS  
 PS Claim 14: 117pp: English.  
 XX  
 XX The sequences given in AAR44516-20 represent mutant versions of the  
 CC adult human alpha globin sequence. These sequences were not given in  
 CC the specification but were generated using the known alpha globin  
 CC sequence. These sequences were used in the production of constructs which  
 CC were used to produce the transgenic pigs of the invention. These  
 CC constructs contain the pig beta globin promoter sequence operably linked  
 CC to the human alpha or beta globin genes. Pigs containing these  
 CC constructs express human haemoglobin (Hb) in their erythrocytes and are  
 CC healthy, suffering no deleterious side effects as a result of  
 CC heterologous Hb production. They can be used as an efficient and  
 CC economical source of human Hb that can be used for transfusion and other  
 CC medical applications. Also, a human/pig hybrid Hb can be produced which  
 CC exhibits a P50 that is higher than that of native human or pig Hb.  
 XX  
 SO Sequence 141 AA.  
 QY 1 PFHLLSHGSAOV 13  
 Db 43 fphdlshsaaay 55  
 Query Match 100.0%; Score 72; DB 15; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 6, 40-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (DNXD-) DNK CORP.  
 XX  
 XX Holzman SH, Kellor H, Kumar P, Logan JS, O'Connell JK;  
 P1 Parsons CT, Pilder SH, Pinkert CA, Sharma A, Swanson ME;  
 P1 White SP;  
 P1 WPI: 1994-007675/01  
 XX  
 XX Transgenic pigs which produce human haemoglobin: conty. DNA encoding  
 P1 human alpha globin and huma beta globin linked to promoters  
 PS  
 PS Claim 14: 117pp: English.  
 XX  
 XX The sequences given in AAR44516-20 represent mutant versions of the  
 CC adult human alpha globin sequence. These sequences were not given in  
 CC the specification but were generated using the known alpha globin  
 CC sequence. These sequences were used in the production of constructs which  
 CC were used to produce the transgenic pigs of the invention. These  
 CC constructs contain the pig beta globin promoter sequence operably linked  
 CC to the human alpha or beta globin genes. Pigs containing these  
 CC constructs express human haemoglobin (Hb) in their erythrocytes and are  
 CC healthy, suffering no deleterious side effects as a result of  
 CC heterologous Hb production. They can be used as an efficient and  
 CC economical source of human Hb that can be used for transfusion and other  
 CC medical applications. Also, a human/pig hybrid Hb can be produced which  
 CC exhibits a P50 that is higher than that of native human or pig Hb.  
 XX  
 SO Sequence 141 AA;  
 QY 1 PFHLLSHGSAOV 13  
 Db 43 fphdlshsaaay 55  
 Query Match 100.0%; Score 72; DB 15; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 6, 40-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

f1 white sh;
f2
f3
f4
f5
f6
f7
f8
f9
fa
fb
fc
fd
fe
ff
fg
fh
fi
fj
fk
fl
fm
fn
fo
fp
fq
fr
fs
ft
fu
fv
fw
fx
fy
fz
ga
gb
gc
gd
ge
gf
gg
gh
gi
gj
gk
gl
gm
gn
go
gp
gq
gr
gs
gt
gu
gv
gw
gx
gy
gz
ha
hb
hc
hd
he
hf
hg
hh
hi
hj
hk
hl
hm
hn
ho
hp
hq
hr
hs
ht
hu
hv
hw
hx
hy
hz
ia
ib
ic
id
ie
if
ig
ih
ii
ij
ik
il
im
in
io
ip
iq
ir
is
it
iu
iv
iv
ix
iy
iz
ja
jb
jc
jd
je
jf
jg
jh
ji
jj
jk
jl
jm
jn
jo
jp
jq
jr
js
jt
ju
jv
jw
jx
jy
jz
ka
kb
kc
kd
ke
kf
kg
kh
ki
kj
kl
km
kn
ko
kp
kq
kr
ks
kt
ku
kv
kw
kx
ky
kz
la
lb
lc
ld
le
lf
lg
lh
li
lj
lk
ll
lm
ln
lo
lp
lq
lr
ls
lt
lu
lv
lw
lx
ly
lz
ma
mb
mc
md
me
mf
mg
mh
mi
mj
mk
ml
mm
mn
mo
mp
mq
mr
ms
mt
mu
mv
mw
mx
my
mz
na
nb
nc
nd
ne
nf
ng
nh
ni
nj
nk
nl
nm
nn
no
np
nq
nr
ns
nt
nu
nv
nw
nx
ny
nz
oa
ob
oc
od
oe
of
og
oh
oi
oj
ok
ol
om
on
oo
op
oq
or
os
ot
ou
ov
ow
ox
oy
oz
pa
pb
pc
pd
pe
pf
pg
ph
pi
pj
pk
pl
pm
pn
po
pp
pq
pr
ps
pt
pu
pv
pw
px
py
pz
qa
qb
qc
qd
qe
qf
qg
qh
qi
qj
qk
ql
qm
qn
qo
qp
qq
qr
qs
qt
qu
qv
qw
qx
qy
qz
ra
rb
rc
rd
re
rf
rg
rh
ri
rj
rk
rl
rm
rn
ro
rp
rq
rr
rs
rt
ru
rv
rw
rx
ry
rz
sa
sb
sc
sd
se
sf
sg
sh
si
sj
sk
sl
sm
sn
so
sp
sq
sr
ss
st
su
sv
sw
sx
sy
sz
ta
tb
tc
td
te
tf
tg
th
ti
tj
tk
tl
tm
tn
to
tp
tq
tr
ts
tt
tu
tv
tw
tx
ty
tz
ua
ub
uc
ud
ue
uf
ug
uh
ui
uj
uk
ul
um
un
uo
up
uq
ur
us
ut
uu
uv
uw
ux
uy
uz
va
vb
vc
vd
ve
vf
vg
vh
vi
vj
vk
vl
vm
vn
vo
vp
vq
vr
vs
vt
vu
vv
vw
vx
vy
vz
wa
wb
wc
wd
we
wf
wg
wh
wi
wj
wk
wl
wm
wn
wo
wp
wq
wr
ws
wt
wu
wv
ww
wx
wy
wz
xa
xb
xc
xd
xe
xf
xg
xh
xi
xj
xk
xl
xm
xn
xo
xp
xq
xr
xs
xt
xu
xv
xw
xx
xy
xz
ya
yb
yc
yd
ye
yf
yg
yh
yi
yj
yk
yl
ym
yn
yo
yp
yq
yr
ys
yt
yu
yv
yw
yx
yy
yz
za
zb
zc
zd
ze
zf
zg
zh
zi
zj
zk
zl
zm
zn
zo
zp
zq
zr
zs
zt
zu
zv
zw
zx
zy
zz

```

```

f1 Transgenic pigs which produce human hemoglobin cont. DNA encoding
f2 human alpha globin and human beta globin linked to promoters
f3
f4
f5
f6
f7
f8
f9
fa
fb
fc
fd
fe
ff
fg
fh
fi
fj
fk
fl
fm
fn
fo
fp
fq
fr
fs
ft
fu
fv
fw
fx
fy
fz
ga
gb
gc
gd
ge
gf
gg
gh
gi
gj
gk
gl
gm
gn
go
gp
gq
gr
gs
gt
gu
gv
gw
gx
gy
gz
ha
hb
hc
hd
he
hf
hg
hh
hi
hj
hk
hl
hm
hn
ho
hp
hq
hr
hs
ht
hu
hv
hw
hx
hy
hz
ia
ib
ic
id
ie
if
ig
ih
ii
ij
ik
il
im
in
io
ip
iq
ir
is
it
iu
iv
iv
ix
iy
iz
ja
jb
jc
jd
je
jf
jg
jh
ji
jj
jk
jl
jm
jn
jo
jp
jq
jr
js
jt
ju
jv
jw
jx
jy
jz
ka
kb
kc
kd
ke
kf
kg
kh
ki
kj
kl
km
kn
ko
kp
kq
kr
ks
kt
ku
kv
kw
kx
ky
kz
la
lb
lc
ld
le
lf
lg
lh
li
lj
lk
ll
lm
ln
lo
lp
lq
lr
ls
lt
lu
lv
lw
lx
ly
lz
ma
mb
mc
md
me
mf
mg
mh
mi
mj
mk
ml
mm
mn
mo
mp
mq
mr
ms
mt
mu
mv
mw
mx
my
mz
na
nb
nc
nd
ne
nf
ng
nh
ni
nj
nk
nl
nm
nn
no
np
nq
nr
ns
nt
nu
nv
nw
nx
ny
nz
oa
ob
oc
od
oe
of
og
oh
oi
oj
ok
ol
om
on
oo
op
oq
or
os
ot
ou
ov
ow
ox
oy
oz
pa
pb
pc
pd
pe
pf
pg
ph
pi
pj
pk
pl
pm
pn
po
pp
pq
pr
ps
pt
pu
pv
pw
px
py
pz
qa
qb
qc
qd
qe
qf
qg
qh
qi
qj
qk
ql
qm
qn
qo
qp
qq
qr
qs
qt
qu
qv
qw
qx
qy
qz
ra
rb
rc
rd
re
rf
rg
rh
ri
rj
rk
rl
rm
rn
ro
rp
rq
rr
rs
rt
ru
rv
rw
rx
ry
rz
sa
sb
sc
sd
se
sf
sg
sh
si
sj
sk
sl
sm
sn
so
sp
sq
sr
ss
st
su
sv
sw
sx
sy
sz
ta
tb
tc
td
te
tf
tg
th
ti
tj
tk
tl
tm
tn
to
tp
tq
tr
ts
tt
tu
tv
tw
tx
ty
tz
ua
ub
uc
ud
ue
uf
ug
uh
ui
uj
uk
ul
um
un
uo
up
uq
ur
us
ut
uu
uv
uw
ux
uy
uz
va
vb
vc
vd
ve
vf
vg
vh
vi
vj
vk
vl
vm
vn
vo
vp
vq
vr
vs
vt
vu
vv
vw
vx
vy
vz
wa
wb
wc
wd
we
wf
wg
wh
wi
wj
wk
wl
wm
wn
wo
wp
wq
wr
ws
wt
wu
wv
ww
wx
wy
wz
xa
xb
xc
xd
xe
xf
xg
xh
xi
xj
xk
xl
xm
xn
xo
xp
xq
xr
xs
xt
xu
xv
xw
xx
xy
xz
ya
yb
yc
yd
ye
yf
yg
yh
yi
yj
yk
yl
ym
yn
yo
yp
yq
yr
ys
yt
yu
yv
yw
yx
yy
yz
za
zb
zc
zd
ze
zf
zg
zh
zi
zj
zk
zl
zm
zn
zo
zp
zq
zr
zs
zt
zu
zv
zw
zx
zy
zz

```

```

Query Match          Score 72 108 152 Length 141:
Best Local Similarity 100.00% Ident. No. 6, 30 00%
Matches 132 Conservative 0% Mismatches 0 Indels 0 Gaps 0
ID 1 PFPHSHSAGV 13
1111111111
ID 43 PPHSHSAGV 55

```

Search completed: June 21, 2001, 10:21:48  
 Job Time: 362 sec



	Sequence	x <sub>i</sub>	Appri
20	72	100.0	US-08-484-341-X
27	72	100.0	US-09-005-546-16

MOLECULAR TYPE: peptide

```
Matches: 13; conservative: 0; mismatches: 0; indels: 0; gaps: 0
```



TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-627-173-2

Query Match: 100.0%; Score 72; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4,8e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFHFDLSHGSAQV 13  
DB 2 PFHFDLSHGSAQV 14

RESULT 5  
US-08-535-882A-2  
Sequence 2: Application US/08/535882A  
Patent No. 5939391  
GENERAL INFORMATION:  
APPLICANT: TSYRLOVA, IRENA  
APPLICANT: WOLFE, STEPHEN D.  
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 NORTH CLARE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPIER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/535,882A  
FILING DATE: 28-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1331-177  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-535-882A-2

Query Match: 100.0%; Score 72; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4,8e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PFHFDLSHGSAQV 13  
DB 2 PFHFDLSHGSAQV 14

RESULT 6  
US-09-005-546-2  
Sequence 2: Application US/09/005546  
Patent No. 6090782  
GENERAL INFORMATION:  
APPLICANT: TSYRLOVA, IRENA  
APPLICANT: WOLFE, STEPHEN D.  
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 NORTH CLARE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201 4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPIER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,546  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/535,882  
FILING DATE: 28-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1331-177  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-005-546-2

Query Match: 100.0%; Score 72; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4,8e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFHFDLSHGSAQV 13  
DB 2 PFHFDLSHGSAQV 14

RESULT 7  
US-07-923-692C-8  
Sequence 8: Application US/07/923692C  
Patent No. 5316931  
GENERAL INFORMATION:  
APPLICANT: Douson, Jon  
APPLICANT: Dawson, William O.  
APPLICANT: Grantham, George L.  
APPLICANT: Turpen, Thomas H.  
APPLICANT: Turpen, Ann Myers  
APPLICANT: Garger, Stephen J.  
APPLICANT: Griffin, Laurence K.  
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Linbach & Linbach  
STREET: 2001 Ferry Building  
CITY: San Francisco

```

1 STATE: CAL
2 ZIP: 94111
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 OPERATING SYSTEM: Macintosh
6 SOFTWARE: Patent in Release #1.00, Version #1.00
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: 08/081700/267
9 FILING DATE: 01 JUN 1992
10 CLASSIFICATION: 435
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: 08/080111
13 FILING DATE: 27 OCT 1990
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: 08/011111
16 FILING DATE: 16 JAN 1991
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 08/010101
19 FILING DATE: 17 FEB 1989
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: 08/100166
22 FILING DATE: 26 FEB 1988
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/100171
25 FILING DATE: 26 FEB 1988
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: 08/047637
28 FILING DATE: 05 MAY 1989
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/000138
31 FILING DATE: 08 JUN 1989
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: 08/019279
34 FILING DATE: 15 JUL 1989
35 ALTERNATIVE/ACUTE INFORMATION:
36 NAME: Hollman, Albert F.
37 REFERENCE NUMBER: 28,702
38 REFERENCE NUMBER: 31-61 20121
39 REFERENCE NUMBER: USA
40 INFORMATIONAL INFORMATION:
41 TELEPHONE: 415 433 4150
42 TELEFAX: 415 433 8716
43 INFORMATION FOR SEQ ID NO: 1:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 141 amino acids
46 TYPE: amino acid
47 SOURCE: human
48 MOLECULE TYPE: protein
49 US 07/063,028 B

```

```

50 Query Match: 100.0% Score: 7.2 E-01 Length: 141
51 Host Local Similarity: 100.0% E-01 No. seq: 092
52 Matches: 132 Conserved: 02 Mismatches: 02 Indels: 02 Gaps: 02

```

```

53 1 PHRED/SHANAW 13
54 43 PHRED/SHANAW 55

```

```

1 CITY: Boulder
2 STATE: Colorado
3 ZIP: 80301
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 OPERATING SYSTEM: Macintosh
7 SOFTWARE: Microsoft Word 6.04
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: 08/081700/0908
10 FILING DATE: December 20, 1993
11 CLASSIFICATION: 540
12 ALTERNATIVE/ACUTE INFORMATION:
13 NAME: No. 55682548K, HENTY, J.
14 REFERENCE NUMBER: 83200
15 REFERENCE NUMBER: 83200
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 303-541 3322
18 TELEFAX: 303-444-3014
19 INFORMATION FOR SEQ ID NO: 1:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 141
22 TYPE: amino acid
23 STRANDEDNESS: single
24 TOPOLOGY: linear
25 MOLECULE TYPE: protein
26 US 08-170 095B 1

```

```

50 Query Match: 100.0% Score: 7.2 E-01 Length: 141
51 Host Local Similarity: 100.0% E-01 No. seq: 092
52 Matches: 132 Conserved: 02 Mismatches: 02 Indels: 02 Gaps: 02

```

```

53 1 PHRED/SHANAW 13
54 43 PHRED/SHANAW 55

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 641,617  
FILING DATE: 16-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 310,881  
FILING DATE: 17-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,766  
FILING DATE: 26-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,771  
FILING DATE: 26-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 347,637  
FILING DATE: 05-MAY-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 363,138  
FILING DATE: 08-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 219,279  
FILING DATE: 19-JUL-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Ballin, Albert P.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: BIOG-20121 USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-443-4150  
TELEFAX: 415-443-8716  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-184-247-8

Query Match 100.0%; Score 72, DB 1, Length 141,  
Best Local Similarity 100.0%; Pred. No. 66-06;  
Matches 13, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 PFFPLSHSGSAV 13  
|||||  
DB 43 PFFPLSHSGSAV 55

RESULT 10  
US-08-240-712-17  
Sequence 17, Application US/08240712  
Patent No. 5599907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STELLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
TITLE OF INVENTION: HEMOGLOBINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Broadway and Melmark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: us-08-240-712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PC/08240712  
FILING DATE: 13-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON 6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Pept130  
US-08-240-712-17

Query Match 100.0%; Score 72, DB 1, Length 141,  
Best Local Similarity 100.0%; Pred. No. 66-06;  
Matches 13, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 PFFPLSHSGSAV 13  
|||||  
DB 43 PFFPLSHSGSAV 55

RESULT 11  
US-08-240-712-26  
Sequence 26, Application US/08240712  
Patent No. 5599907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STELLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
TITLE OF INVENTION: HEMOGLOBINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Broadway and Melmark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: us-08-240-712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PC/08240712  
FILING DATE: 13-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 248633  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single



SOFTWARE: Patentio Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/443,890  
 FILING DATE: 31-MAY-1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/240,712  
 FILING DATE: 09-MAY-1994  
 APPLICATION NUMBER: PCT/US92/09752  
 FILING DATE: 14-MAY-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: COOPER, IVER P  
 REGISTRATION NUMBER: 28,005  
 REFERENCE/DCKET NUMBER: ANDERSON-6  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-628-5197  
 TELEFAX: 202-747-3528  
 TELEX: 248633  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 141 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-443-890-17

Query Match 100.0% Score 72; FR 1; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 60-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFPHDSHSAQV 13  
 DB 43 PFPHDSHSAQV 55

RESULT 15  
 US-08-443-890-26  
 Sequence 26, Application US/08443890  
 Patent No. 5739011  
 GENERAL INFORMATION:  
 APPLICANT: ANDERSON, DAVID C.  
 APPLICANT: MAHEWS, ANTONY JAMES  
 APPLICANT: STETLER, GARY L.  
 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Broadway and Neimark  
 STREET: 419 Seventh Street, N.W., Suite 300  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentio Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/443,890  
 FILING DATE: 31-MAY-1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/240,712  
 FILING DATE: 09-MAY-1994  
 APPLICATION NUMBER: PCT/US92/09752  
 FILING DATE: 14-MAY-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: COOPER, IVER P  
 REGISTRATION NUMBER: 28,005  
 REFERENCE/DCKET NUMBER: ANDERSON-6

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-628-5197  
 TELEFAX: 202-747-3528  
 TELEX: 248633  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 141 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-443-890-26

Query Match 100.0% Score 72; FR 1; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 60-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFPHDSHSAQV 13  
 DB 43 PFPHDSHSAQV 55

Search completed: June 21, 2001, 10:25:09  
 Job Time: 258 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2001, 10:24:21 Search time: 20.16 seconds  
(without alignments)  
49,120 Million cell updates/sec

Title: US-08-832-443b-1

72

Perfect score: 1 FPHFDLSHGSACV 13

Scoring table: BlastDM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

pk\_68:\*  
1: pfr1:\*  
2: pfr2:\*  
3: pfr3:\*  
4: pfr4:\*

Pred. No. is the number of positive predictions by above, to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	100.0	140	1 HABOKA	hemoglobin alpha c
2	72	100.0	141	1 HAOCP	hemoglobin alpha c
3	72	100.0	141	1 HAOCP	hemoglobin alpha c
4	72	100.0	141	1 HAOCP	hemoglobin alpha c
5	72	100.0	141	1 HAOCP	hemoglobin alpha c
6	72	100.0	141	1 HAOCP	hemoglobin alpha c
7	72	100.0	141	1 HAOCP	hemoglobin alpha c
8	72	100.0	141	1 HAOCP	hemoglobin alpha c
9	72	100.0	141	1 HAOCP	hemoglobin alpha c
10	72	100.0	141	1 HAOCP	hemoglobin alpha c
11	72	100.0	141	1 HAOCP	hemoglobin alpha c
12	72	100.0	141	1 HAOCP	hemoglobin alpha c
13	72	100.0	141	1 HAOCP	hemoglobin alpha c
14	72	100.0	141	1 HAOCP	hemoglobin alpha c
15	72	100.0	141	1 HAOCP	hemoglobin alpha c
16	72	100.0	141	1 HAOCP	hemoglobin alpha c
17	72	100.0	141	1 HAOCP	hemoglobin alpha c
18	72	100.0	141	1 HAOCP	hemoglobin alpha c
19	72	100.0	141	1 HAOCP	hemoglobin alpha c
20	72	100.0	141	1 HAOCP	hemoglobin alpha c
21	72	100.0	141	1 HAOCP	hemoglobin alpha c
22	72	100.0	141	1 HAOCP	hemoglobin alpha c
23	72	100.0	141	1 HAOCP	hemoglobin alpha c
24	72	100.0	141	1 HAOCP	hemoglobin alpha c
25	72	100.0	141	1 HAOCP	hemoglobin alpha c
26	72	100.0	141	1 HAOCP	hemoglobin alpha c
27	72	100.0	141	1 HAOCP	hemoglobin alpha c
28	72	100.0	141	1 HAOCP	hemoglobin alpha c
29	72	100.0	141	1 HAOCP	hemoglobin alpha c

30	72	100.0	141	1 HAOCP	hemoglobin alpha c
31	72	100.0	141	1 HAOCP	hemoglobin alpha c
32	72	100.0	141	1 HAOCP	hemoglobin alpha c
33	72	100.0	141	1 HAOCP	hemoglobin alpha c
34	72	100.0	141	1 HAOCP	hemoglobin alpha c
35	72	100.0	141	1 HAOCP	hemoglobin alpha c
36	72	100.0	141	1 HAOCP	hemoglobin alpha c
37	72	100.0	141	1 HAOCP	hemoglobin alpha c
38	72	100.0	141	1 HAOCP	hemoglobin alpha c
39	72	100.0	141	1 HAOCP	hemoglobin alpha c
40	72	100.0	141	1 HAOCP	hemoglobin alpha c
41	72	100.0	141	1 HAOCP	hemoglobin alpha c
42	72	100.0	141	1 HAOCP	hemoglobin alpha c
43	72	100.0	141	1 HAOCP	hemoglobin alpha c
44	72	100.0	141	1 HAOCP	hemoglobin alpha c
45	72	100.0	141	2 S20278	hemoglobin alpha c

#### ALIGNMENTS

RESULT 1  
HABOKA  
hemoglobin alpha chain - greater kudu  
C:Species: Tragelaphus streptoceros (greater kudu)  
C:Date: 04-Nov-1986 #Sequence\_revision 04-Nov-1986 #Text\_change 03-Mar-2000  
C:Accession: A02293  
R:Rotewald, K.; Wiesner, H.; Braunitzer, G.  
Biol. Chem. Hoppe-Seyler 366, 395-402, 1985  
A:Title: Primary structure of the hemoglobins from the greater kudu antelope (Tragelaphus streptoceros)  
A:Reference number: A00684; MUID:85279893  
A:Accession: A02293  
A:Molecule type: protein  
A:Residues: 1-140 <8CD>  
C:Superfamily: globin hemoglobin  
C:Keywords: blood, erythrocyte, heme, iron, metal, protein, oxygen car  
E:1-140/Domain: globin hemoglobin <8CD>  
F:57/binding site: alpha (His) (distal axial ligand) #status predicted  
F:86/binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 140;  
Best local similarity 100.0%; Pred. No. 8; Le-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 42 FPHFDLSHGSACV 54  
RESULT 2  
HABOKA  
hemoglobin alpha chain - pygmy chimpanzee  
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
C:Date: 30-Jun-1993 #Sequence\_revision 30-Jun-1993 #Text\_change 03-Mar-2000  
C:Accession: C94303  
R:Goodman, M.; Braunitzer, G.; Stangl, A.; Schrank, B.  
Nature 303, 546-548, 1983  
A:Title: Evidence on human origins from haemoglobins of African apes.  
A:Reference number: A93303; MUID:84219265  
A:Accession: C94303  
A:Molecule type: protein  
A:Residues: 1-141 <8CD>  
C:Superfamily: globin hemoglobin  
C:Keywords: blood, erythrocyte, heme, iron, metalloprotein, oxygen car  
E:57/binding site: alpha (His) (distal axial ligand) #status predicted  
E:58/binding site: alpha (His) (proximal axial ligand) #status predicted  
F:87/binding site: heme iron (His) (proximal axial ligand) #status predicted  
Query Match 100.0%; Score 72; DB 1; Length 141;  
Best local similarity 100.0%; Pred. No. 9; Le-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



R:Mahoney, W.C.: Nure, P.E.

Arch. Biochem. Biophys. 196, 64-72, 1979

A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(I

A:Reference number: S06514; MUID:80063760

A:Accession: S06514

A:Molecule type: Protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

RESULT 8

HAMQZP hemoglobin alpha-11 chain - pig-tailed macaque

C:Species: Macaca nemestrina (pig-tailed macaque)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 03-Mar-2000

C:Accession: S07670

R:Mahoney, W.C.: Nure, P.E.

Arch. Biochem. Biophys. 196, 64-72, 1979

A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(I

A:Reference number: S06514; MUID:80063760

A:Accession: S07670

A:Molecule type: protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

RESULT 9

HAMKP hemoglobin alpha chain - black-handed spider monkey

C:Species: Atelos geoffroyi (black-handed spider monkey)

C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 03-Mar-2000

C:Accession: A02253

R:Matsumoto, G.: Maita, T.: Sugiyama, Y.: Saitoh, M.: Ota, Y.: Ariga, A.: Goodman, M.

Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973

A:Title: Studies on the primary structures of alpha and beta polypeptide chains of adult

A:Reference number: A02253; MUID:75059125

A:Accession: A02253

A:Molecule type: protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

RESULT 10

HACCB hemoglobin alpha chain - black-tailed marmoset

C:Species: Callithrix jacchus (black-tailed marmoset)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 03-Mar-2000

C:Accession: S06512

R:Matsumoto, G.: Maita, T.: Sugiyama, Y.: Saitoh, M.: Ota, Y.: Ariga, A.: Goodman, M.

Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973

A:Title: Studies on the primary structures of alpha and beta polypeptide chains of adult h

A:Reference number: A02254; MUID:75059125

A:Accession: A02254

A:Molecule type: protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen car

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

RESULT 11

HAMQA hemoglobin alpha chain - brown capuchin

C:Species: Cebus apella (brown capuchin, black capped capuchin)

C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 03-Mar-2000

C:Accession: A02254

R:Matsumoto, G.: Maita, T.: Sugiyama, Y.: Saitoh, M.: Ota, Y.: Ariga, A.: Goodman, M.

Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973

A:Title: The amino acid sequences of the alpha and beta polypeptide chains of adult h

A:Reference number: A02254; MUID:75059125

A:Accession: A02254

A:Molecule type: protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen car

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

RESULT 12

HAMQF hemoglobin alpha chain - brown-headed tamarin

C:Species: Saguinus fuscicollis (brown-headed tamarin, saddle-backed tamarin)

C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 03-Mar-2000

C:Accession: A02255

R:Lin, K.D.: Kim, Y.K.: Chernoff, A.L.

Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973

A:Title: Studies on the primary structures of alpha and beta polypeptide chains of adult h

A:Reference number: A02255; MUID:75059125

A:Accession: A02255

A:Molecule type: protein

A:Residues: 1-141 <MA>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen car

F:2-141/domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0% Score 72; DB 1; Length 141;

Best Local Similarity 100.0% Pred. No. 8,2e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLHSGSAOV 13

DB 43 FPHFDLHSGSAOV 55

hemoglobin, sheet: 24, 477-480, 1976

Active: Primary structure of the hemo globin (various isoforms) hemoglobin, 1, 1980 of

Accession number: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

best local similarity: 100.0% score: 7.2, 10.1, length: 141  
 matches: 132, conservation: 0, mismatches: 0, indels: 0, gaps: 0

1 PPHHSHSAAV 13

1 PPHHSHSAAV 55

15

hemoglobin alpha chain - slow form

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

Accession: A00226

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuLink Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2001, 10:25:11 : Search time 11.96 Seconds

(without alignments)  
37,234 Million cell updates/sec

Title: US-08-832-443b-1

Sequence: 1 FPHFDLSHSAQV 13

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 9435 seqs, 3425546 residues

Total number of hits satisfying chosen parameters: 9435

Minimum DB seq length: 0  
Maximum DB seq length: 20000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	100.0	140	1 HBA_TFAST	P04237 Trachelaphus
2	72	100.0	141	1 HBA_TADAR	P11755 tadarda br
3	72	100.0	141	1 HBA2_ARCGA	P41330 arctoccephal
4	72	100.0	141	1 HBA2_HUSMT	P01968 bos mutus g
5	72	100.0	141	1 HBA_ARCGA	P01971 alces alces
6	72	100.0	141	1 HBA_ANTPA	P14387 antilocapra p
7	72	100.0	141	1 HBA_ATECE	P01927 alces alces
8	72	100.0	141	1 HBA_BISBO	P04423 bischof katus
9	72	100.0	141	1 HBA_BOVIN	P01966 bos taurus
10	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
11	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
12	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
13	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
14	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
15	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
16	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
17	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
18	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
19	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
20	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
21	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
22	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
23	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
24	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
25	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
26	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
27	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
28	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
29	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
30	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
31	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
32	72	100.0	141	1 HBA_BISBO	P01966 bos taurus
33	72	100.0	141	1 HBA_BISBO	P01966 bos taurus

34	72	100.0	141	1 HBA_MACAS	P21766 macaca assa
35	72	100.0	141	1 HBA_MACAS	P21767 macaca fass
36	72	100.0	141	1 HBA_MACAS	P24659 macaca fass
37	72	100.0	141	1 HBA_MACAS	P24659 macaca fass
38	72	100.0	141	1 HBA_MACAS	P21902 macaca fass
39	72	100.0	141	1 HBA_MACAS	P21768 macaca fass
40	72	100.0	141	1 HBA_MACAS	P07402 macaca fass
41	72	100.0	141	1 HBA_MACAS	P11751 macaca fass
42	72	100.0	141	1 HBA_MACAS	P11751 macaca fass
43	72	100.0	141	1 HBA_MACAS	P24600 macaca fass
44	72	100.0	141	1 HBA_MACAS	P20243 macaca fass
45	72	100.0	141	1 HBA_MACAS	P24601 macaca fass

## ALIGNMENTS

RESULT 1	HBA_TFAST	STANDARD	PRT: 140 AA.
ID	HBA_TFAST		
AC	P04237		
DT	20-MAR-1987 (Rel. 04, Last sequence update)		
DT	20-MAR-1987 (Rel. 04, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	HEMOLYTHIN ALPHA CHAIN.		
OS	Trachelaphus streptoceros (Greater kudu).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eulalia; Cetartiodactyla; Fungimammalia; Fungimammalia; Fungimammalia;		
OC	Bovidae; Bovinae; Tragelaphus.		
OX	NCBI_TaxID=9446;		
RN	[1]		
RP	SEQUENCE.		
RY	MEDLINE: 85279693; PubMed 4026993;		
KA	Rodwald K., Wiesner H., Braunitzer G.;		
RT	"Primary structure of the hemoglobins from the greater kudu antelope		
RT	(Trachelaphus streptoceros)."		
RL	Biol. Chem. Heppe-Seyler 366:395-402(1985).		
PIR	PIR: A02293; HAHK.A.		
DR	HSP: P01966; HBA.		
DR	InterPro: IPR000471;		
DR	InterPro: IPR002438;		
DR	PRIM: PR00042; Globin, 1.		
DR	PRIM: PR00012; ALPHABEM.		
DR	PROSITE: PS01033; GLOBIN, 1.		
KM	Swiss: 05994 57 57 160N (HEME DISTAL LIGAND).		
PT	METAL.		
FT	METAL.		
SO	SEQUENCE 140 AA. 1449 MK. A401FA49F4F744 C8C64;		
Query Match	100.0%; Score 72; DB 1; Length 140;		
Foot print	Similarity 100.0%; Pct 2.8e-06;		
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 FPHFDLSHSAQV 13		
DB	1 FPHFDLSHSAQV 54		
RESULT 2	HBA_TADAR	STANDARD	PRT: 141 AA.
ID	HBA_TADAR		
AC	P11755		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	HEMOLYTHIN ALPHA-1 CHAIN.		
OS	Tadarda brasiliensis (Brazilian tree-tailed bat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eulalia; Chiroptera; Microchiroptera; Molossid;		
OC	Tadarida.		
OX	NCBI_TaxID=9438;		
RN	[1]		





	2007	Matches	100.0%	Scores	7-7	0% 1-	Length	141-
	Best	Local	Stim	Lut	117	100.0%	Freq.	No. 2, 98; 0%
	Matches	14-	Concavatio	0-	Mismatches	0-	Indels	0-
55	1	FFHHHSHHSNAY	14-					
		111111111111						
68	44	FFHHHSHHSNAY	14-					

REF ID:	9	STANDARD:	PRC:	141 AA.
10	HA POSTH			
11	HA POSTH			
12	HA POSTH			
13	HA POSTH			
14	HA POSTH			
15	HA POSTH			
16	HA POSTH			
17	HA POSTH			
18	HA POSTH			
19	HA POSTH			
20	HA POSTH			
21	HA POSTH			
22	HA POSTH			
23	HA POSTH			
24	HA POSTH			
25	HA POSTH			
26	HA POSTH			
27	HA POSTH			
28	HA POSTH			
29	HA POSTH			
30	HA POSTH			
31	HA POSTH			
32	HA POSTH			
33	HA POSTH			
34	HA POSTH			
35	HA POSTH			
36	HA POSTH			
37	HA POSTH			
38	HA POSTH			
39	HA POSTH			
40	HA POSTH			
41	HA POSTH			
42	HA POSTH			
43	HA POSTH			
44	HA POSTH			
45	HA POSTH			
46	HA POSTH			
47	HA POSTH			
48	HA POSTH			
49	HA POSTH			
50	HA POSTH			
51	HA POSTH			
52	HA POSTH			
53	HA POSTH			
54	HA POSTH			
55	HA POSTH			
56	HA POSTH			
57	HA POSTH			
58	HA POSTH			
59	HA POSTH			
60	HA POSTH			
61	HA POSTH			
62	HA POSTH			
63	HA POSTH			
64	HA POSTH			
65	HA POSTH			
66	HA POSTH			
67	HA POSTH			
68	HA POSTH			
69	HA POSTH			
70	HA POSTH			
71	HA POSTH			
72	HA POSTH			
73	HA POSTH			
74	HA POSTH			
75	HA POSTH			
76	HA POSTH			
77	HA POSTH			
78	HA POSTH			
79	HA POSTH			
80	HA POSTH			
81	HA POSTH			
82	HA POSTH			
83	HA POSTH			
84	HA POSTH			
85	HA POSTH			
86	HA POSTH			
87	HA POSTH			
88	HA POSTH			
89	HA POSTH			
90	HA POSTH			
91	HA POSTH			
92	HA POSTH			
93	HA POSTH			
94	HA POSTH			
95	HA POSTH			
96	HA POSTH			
97	HA POSTH			
98	HA POSTH			
99	HA POSTH			
100	HA POSTH			

RESULT	10
HRA_RECVIN	
11. HRA_LEAVEIN	STANDARD: FRT: 14. AM.
AV	POL966
12	21-JUL-1986 (Rel. 01, 'Treated')
13	21-JUL-1986 (Rel. 01, 'Last sequence update')
14	01-OCT-2000 (Rel. 40, 'Last annotation update')
15	HEMOGLOBIN ALPHA CHAIN.
16	HRA.
17	Bos taurus (bovine).
18	ElkayAfor: Metazoa: Chordata: Vertebrata: Euteleostomi:
19	Mammalia: Eutheria: Cetartiodactyla: Eumastixia: Bovidae:
20	Bovidae: Bovinae: Bos.
21	NCBI TaxID 99141
22	[1]

FA	POTURY M.F., Form G5; Boyant et al.; Kentucky J. Science 17:
GA	"A novel allosteric mechanism in haemoglobin: Structure of bovine
RT	deoxyhaemoglobin: absence of specific chloride binding sites and
RT	location of the chloride-linked holo effect in bovine and human
R1	haemoglobin."
R1	J. Mol. Biol. 233:53-64(1993).
CY	-1- PUNJ LONE INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CY	VARIOUS PERIPHERAL TISSUES.
CY	-1- SUBUNIT: Tetramer of two ALPHA CHAINS AND TWO BETA CHAINS.
CY	-1- TISSUE SPECIFICITY: RBC BLOOD CELLS.
CY	-1- SIMILARITY: BELONGS TO THE Globin FAMILY.
CY	-1- DATABASE: NAME=worthington enzyme manual;
	WWW.WORTHINGTON.CCNC.BIOLOGY.DUKE.EDU/ENZYME/HB.html "
PR	PR: AUZ289; HMO.
DB	PRO: UNP: Q1 MAY 94.
DB	InterPro: IPRO00971; -1-
DB	Td cPiro: IPRO02498; -1-
DB	Trame: IPRO0442; globin 1.
DB	PRINTS: PRO0612; ALPHAHBM.
DB	PROSITE: PS01043; GLOBIN_1.
KW	Heme; oxygen transport; Respiratory protein; Erythrocyte
	disstructure.
F1	METAL 50 53 15-N (HBM METAL LIGAND)
F1	METAL 67 67 15-N (HBM PROXIMAL LIGAND).
S2	SEQUENCE 141 AA; 150's MW: 4100AA+100'DNA +6'94'

```

Query Match: Score: 72: 401: Length: 141:
Post_Local_Similarity: 100.00: Pctid: No. 2: 60:00:
Matches: 142: Conserved: 02: Missed: 02: Gaps: 02:
Q7: 1 PROPHETIN,SAUV 14
|||||
DB: 43 PROPHETIN,SAUV 55

RESULT: 11
HBA_CATAL: STANDARD: PP: 141 AA.
Ac: 139/22

```



DI 01-NOV-1999 (rel. 16, created)  
 DT 01-NOV-1990 (rel. 16, last sequence update)  
 DI 01-OCT-2000 (rel. 46, last annotation update)  
 DE HEMOGLOBIN ALPHA CHAIN.  
 OS Callithrix argentatus (Black-tailed marmoset).  
 OC Pukaryota; Metazoa; Chordata; Gracilaria; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;  
 CC Callitrich.  
 NC NCBI\_TaxID=9482;  
 RN 11  
 RP SEQUENCE.  
 RX MEDLINE=8421238; PubMed=5127202;  
 RA Maita F., Hayashida M., Matsuda G.;  
 RT "Primary structures of adult hemoglobins of silvery marmoset,  
 Callithrix argentatus, and cotton-tailed marmoset, Leontideus rosalia";  
 RL J. Biochem. 55: 805-813(1984)  
 CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE  
 CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.  
 CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS  
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY  
 DR PIR: S06512; HMOJB.  
 DR HSSP: P01922; 1FGB.  
 DR InterPro: IPR000971;  
 DR InterPro: IPR002338;  
 DR Pfam: PF00042; globin; 1.  
 DR PRINTS: PR00612; ALPHAHAM.  
 DR PROSITE: PS01033; GLOBIN; 1.  
 KW Heme; Oxygen transport; Respiratory protein; Erythrocyte.  
 FT METAL 58  
 FT METAL 87  
 SU SEQUENCE 141 AA; 15138 MW; 26B13C7E2A68BC02 CRO64;  
 QY 1 PFHFDLSHNSAOV 13  
 DB 43 PFHFDLSHNSAOV 55  
 RESULT 12  
 ID HBA\_CAMPB STANDARD; PRI: 141 AA.  
 AC P01974;  
 DT 21-JUL-1986 (rel. 01, created)  
 DT 21-JUL-1986 (rel. 01, last sequence update)  
 DT 01-OCT-2000 (rel. 40, last annotation update)  
 DE HEMOGLOBIN ALPHA CHAIN.  
 OS Camelus dromedarius (Dromedary) (Arabian camel), and  
 OS Camelus bactrianus (Bactrian camel).  
 OC Eukaryota; Metazoa; Chordata; Gracilaria; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 NC NCBI\_TaxID=9839; 9837;  
 RN 11  
 RP SEQUENCE.  
 RX SPECIES=C.dromedarius;  
 RX MEDLINE=80114125; PubMed=527943;  
 RA Ramanitzy G., Schrank R., Stangl A., Wiesner H.;  
 RT "Respiration at high altitudes: phosphate-protein interaction: the  
 RT sequence of hemoglobins from quinea pig and dromedary";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1941-1946(1979).  
 RN 12  
 RP SEQUENCE.  
 RX SPECIES=C.bactrianus;  
 RA Ramanitzy G., Schrank R., Stangl A., Wiesner H.;  
 RT "Respiration at high altitudes: phosphate-protein interaction:  
 RT sequence of the hemoglobins of the hamster (Mesocricetus auratus) and  
 RT the camel (Camelus ferus, Camelidae).";  
 RL J. Chem. Soc. Pak 2:1-7(1980)  
 CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.  
 CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.  
 DR PIR: B92758; HMOJA.  
 DR PIR: B91685; HMOJA.  
 DR HSSP: P01965; 2FGB.  
 DR InterPro: IPR000971;  
 DR InterPro: IPR002338;  
 DR Pfam: PF00042; globin; 1.  
 DR PRINTS: PR00612; ALPHAHAM.  
 DR PROSITE: PS01033; GLOBIN; 1.  
 KW Heme; Oxygen transport; Respiratory protein; Erythrocyte.  
 FT METAL 58  
 FT METAL 87  
 SU SEQUENCE 141 AA; 15172 MW; 26B06F6B676E50 CRO64;  
 QY 1 PFHFDLSHNSAOV 13  
 DB 43 PFHFDLSHNSAOV 55  
 RESULT 13  
 ID HBA\_CAPHI STANDARD; PRI: 141 AA.  
 AC P01970;  
 DT 21-JUL-1986 (rel. 01, created)  
 DT 21-JUL-1986 (rel. 01, last sequence update)  
 DT 30-MAY-2000 (rel. 39, last annotation update)  
 DE HEMOGLOBIN ALPHA-1 AND ALPHA-2 CHAINS.  
 OS Capra hircus (Goat).  
 OS Capra aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Gracilaria; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 NC NCBI\_TaxID=9925; 9899; 9940;  
 RN 11  
 RP SEQUENCE FROM N.A. (ALPHA 1 AND 2).  
 RX SPECIES=C.hircus;  
 RX MEDLINE=82213618; PubMed=6382925;  
 RA Sehn E.A., Wernke S.M., Linarel J.B.;  
 RT "Gene conversion of two functional goat alpha globin genes preserves  
 RT only minimal flanking sequences";  
 RL J. Biol. Chem. 267:6825-6835(1982).  
 RN 12  
 RP PARTIAL SEQUENCE.  
 RX SPECIES=C.hircus;  
 RX MEDLINE=68113123; PubMed=5658544;  
 RA Huisman T.H.J., Brandt G., Wilson J.B.;  
 RT "The structure of goat hemoglobins. II. Structural studies of the  
 RT alpha chains of the hemoglobins A and B";  
 RL J. Biol. Chem. 243:3675-3686(1968).  
 RN 13  
 RP SEQUENCE.  
 RX SPECIES=A.tervial;  
 RX MEDLINE=70162483; PubMed=5439729;  
 RA Wilson J.B., Wightstone K.N., Huisman T.H.J.;  
 RT "Hemoglobin alpha chain deletion in Barbary sheep, A.tervial;  
 RT tervial; Pallas; 1777";  
 RL Nature 226:354-355(1970).  
 RN 14  
 RP SEQUENCE.  
 RX SPECIES=Sheep; STEVAIN; GOAT; BARBED; AND; GOAT; HBB30;  
 RX MEDLINE=67209244; PubMed=6033754;  
 RA Beale D.;  
 RT "A partial amino acid sequence for sheep haemoglobin A.";  
 RL Biochem. J. 133:129-140(1967).  
 RN 15

SEQUENCE.

67 SPECIES SHEEP STRAIN RAMBOUTLET BREEDS

68 MEDLINE 664124 PubMed 649645

69 WILSON I H, Brand C, Huxham I H, et al

70 "The structure of sheep hemoglobins: A structural studies on the

71 alpha chain of hemoglobin A."

72 J Mol Chem 24:6007-6012(1968)

73

74 VARIANT ALLELES

75

76 SPECIES SHEEP

77 MEDLINE 664124 PubMed 649645

78 WILSON I H, Brand C, Huxham I H, et al

79 "The structure of sheep hemoglobins: A structural studies on the

80 alpha chain of hemoglobin A."

81 J Mol Chem 24:6007-6012(1968)

82

83 VARIANT ALLELES

84

85 SPECIES SHEEP

86 MEDLINE 664124 PubMed 649645

87 WILSON I H, Brand C, Huxham I H, et al

88 "The structure of sheep hemoglobins: A structural studies on the

89 alpha chain of hemoglobin A."

90 J Mol Chem 24:6007-6012(1968)

91

92 SPECIES SHEEP

93 MEDLINE 664124 PubMed 649645

94 WILSON I H, Brand C, Huxham I H, et al

95 "The structure of sheep hemoglobins: A structural studies on the

96 alpha chain of hemoglobin A."

97 J Mol Chem 24:6007-6012(1968)

98

99 SPECIES SHEEP

100 MEDLINE 664124 PubMed 649645

101 WILSON I H, Brand C, Huxham I H, et al

102 "The structure of sheep hemoglobins: A structural studies on the

103 alpha chain of hemoglobin A."

104 J Mol Chem 24:6007-6012(1968)

105

106 SPECIES SHEEP

107 MEDLINE 664124 PubMed 649645

108 WILSON I H, Brand C, Huxham I H, et al

109 "The structure of sheep hemoglobins: A structural studies on the

110 alpha chain of hemoglobin A."

111 J Mol Chem 24:6007-6012(1968)

112

113 SPECIES SHEEP

114 MEDLINE 664124 PubMed 649645

115 WILSON I H, Brand C, Huxham I H, et al

116 "The structure of sheep hemoglobins: A structural studies on the

117 alpha chain of hemoglobin A."

118 J Mol Chem 24:6007-6012(1968)

119

120 SPECIES SHEEP

121 MEDLINE 664124 PubMed 649645

122 WILSON I H, Brand C, Huxham I H, et al

123 "The structure of sheep hemoglobins: A structural studies on the

124 alpha chain of hemoglobin A."

125 J Mol Chem 24:6007-6012(1968)

126

127 SPECIES SHEEP

128 MEDLINE 664124 PubMed 649645

129 WILSON I H, Brand C, Huxham I H, et al

130 "The structure of sheep hemoglobins: A structural studies on the

131 alpha chain of hemoglobin A."

132 J Mol Chem 24:6007-6012(1968)

133

134 SPECIES SHEEP

135 MEDLINE 664124 PubMed 649645

136 WILSON I H, Brand C, Huxham I H, et al

137 "The structure of sheep hemoglobins: A structural studies on the

138 alpha chain of hemoglobin A."

139 J Mol Chem 24:6007-6012(1968)

140

141 SPECIES SHEEP

142 MEDLINE 664124 PubMed 649645

143 WILSON I H, Brand C, Huxham I H, et al

144 "The structure of sheep hemoglobins: A structural studies on the

145 alpha chain of hemoglobin A."

146 J Mol Chem 24:6007-6012(1968)

147

148 SPECIES SHEEP

149 MEDLINE 664124 PubMed 649645

150 WILSON I H, Brand C, Huxham I H, et al

151 "The structure of sheep hemoglobins: A structural studies on the

152 alpha chain of hemoglobin A."

153 J Mol Chem 24:6007-6012(1968)

154

155 SPECIES SHEEP

156 MEDLINE 664124 PubMed 649645

157 WILSON I H, Brand C, Huxham I H, et al

158 "The structure of sheep hemoglobins: A structural studies on the

159 alpha chain of hemoglobin A."

160 J Mol Chem 24:6007-6012(1968)

161

162 SPECIES SHEEP

163 MEDLINE 664124 PubMed 649645

164 WILSON I H, Brand C, Huxham I H, et al

165 "The structure of sheep hemoglobins: A structural studies on the

166 alpha chain of hemoglobin A."

167 J Mol Chem 24:6007-6012(1968)

168

169 SPECIES SHEEP

170 MEDLINE 664124 PubMed 649645

171 WILSON I H, Brand C, Huxham I H, et al

172 "The structure of sheep hemoglobins: A structural studies on the

173 alpha chain of hemoglobin A."

174 J Mol Chem 24:6007-6012(1968)

175

176 SPECIES SHEEP

177 MEDLINE 664124 PubMed 649645

178 WILSON I H, Brand C, Huxham I H, et al

179 "The structure of sheep hemoglobins: A structural studies on the

180 alpha chain of hemoglobin A."

181 J Mol Chem 24:6007-6012(1968)

182

183 SPECIES SHEEP

184 MEDLINE 664124 PubMed 649645

185 WILSON I H, Brand C, Huxham I H, et al

186 "The structure of sheep hemoglobins: A structural studies on the

187 alpha chain of hemoglobin A."

188 J Mol Chem 24:6007-6012(1968)

189

190 SPECIES SHEEP

191 MEDLINE 664124 PubMed 649645

192 WILSON I H, Brand C, Huxham I H, et al

193 "The structure of sheep hemoglobins: A structural studies on the

194 alpha chain of hemoglobin A."

195 J Mol Chem 24:6007-6012(1968)

196

197 SPECIES SHEEP

198 MEDLINE 664124 PubMed 649645

199 WILSON I H, Brand C, Huxham I H, et al

200 "The structure of sheep hemoglobins: A structural studies on the

201 alpha chain of hemoglobin A."

202 J Mol Chem 24:6007-6012(1968)

203

204 SPECIES SHEEP

205 MEDLINE 664124 PubMed 649645

206 WILSON I H, Brand C, Huxham I H, et al

207 "The structure of sheep hemoglobins: A structural studies on the

208 alpha chain of hemoglobin A."

209 J Mol Chem 24:6007-6012(1968)

210

211 SPECIES SHEEP

212 MEDLINE 664124 PubMed 649645

213 WILSON I H, Brand C, Huxham I H, et al

214 "The structure of sheep hemoglobins: A structural studies on the

215 alpha chain of hemoglobin A."

216 J Mol Chem 24:6007-6012(1968)

217

218 SPECIES SHEEP

219 MEDLINE 664124 PubMed 649645

220 WILSON I H, Brand C, Huxham I H, et al

221 "The structure of sheep hemoglobins: A structural studies on the

222 alpha chain of hemoglobin A."

223 J Mol Chem 24:6007-6012(1968)

224

225 SPECIES SHEEP

226 MEDLINE 664124 PubMed 649645

227 WILSON I H, Brand C, Huxham I H, et al

228 "The structure of sheep hemoglobins: A structural studies on the

229 alpha chain of hemoglobin A."

230 J Mol Chem 24:6007-6012(1968)

Query Match: Locals Score 7.2 E-11 Length 141

Best Local Similarity: 100.0% Prod. No. 2 Mo 066

Matches: 13 Conserved Pts: 0 Mismatches: 0 Indels: 0 Gaps: 0

1. PHEIDISBNAV 14

43 PHEIDISBNAV 55

RESULT 14

10 HMA\_CERVA STANDARD: PRT: 141 AA.

11 HMA\_CERVA

12 P01928:

13 21-JUN-1986 (Ref: 01, Created)

14 21-JUN-1986 (Ref: 01, Last sequence update)

15 01-JUN-1994 (Ref: 29, Last annotation update)

16 HMA\_CERVA ALPHA CHAIN.

17 Cebus apella (brown capped capuchin).

18 Eukaryotic Metazoan Chordata: Vertebrate: Eucaryotom:

19 Mammalia: Eutheria: Primates: Platyrrhini: Cebidae: Cebus.

20 NCRL\_T0010 95157

21

22 SPECIES

23 MEDLINE 75059125 PubMed 4215719

24 Matsuda G, Mita T, Watanabe K, Ona Y,

25 Goodman M, Barnabas J, Pychko W,

26 "The amino acid sequences of the alpha and beta polypeptide chains of

27 adult hemoglobin of the capuchin monkey (Cebus apella)."?

28 J Hematol 150:157-161 (1973)

29

30 PIR: A02294: HMA\_CERVA

31 HSSP: P01922: PHEIDISBNAV

32 InterPro: IP000971:

33 InterPro: IP000971:

34 PIR: A02294: HMA\_CERVA

35 PIR: A02294: HMA\_CERVA

36 PIR: A02294: HMA\_CERVA

37 HMA\_CERVA ALPHA CHAIN.

38 HMA\_CERVA ALPHA CHAIN.

39 HMA\_CERVA ALPHA CHAIN.

40 HMA\_CERVA ALPHA CHAIN.

41 HMA\_CERVA ALPHA CHAIN.

42 HMA\_CERVA ALPHA CHAIN.

43 HMA\_CERVA ALPHA CHAIN.

44 HMA\_CERVA ALPHA CHAIN.

45 HMA\_CERVA ALPHA CHAIN.

46 HMA\_CERVA ALPHA CHAIN.

47 HMA\_CERVA ALPHA CHAIN.

48 HMA\_CERVA ALPHA CHAIN.

49 HMA\_CERVA ALPHA CHAIN.

50 HMA\_CERVA ALPHA CHAIN.

Query Match: Locals Score 7.2 E-11 Length 141

Best Local Similarity: 100.0% Prod. No. 2 Mo 066

Matches: 13 Conserved Pts: 0 Mismatches: 0 Indels: 0 Gaps: 0

1. PHEIDISBNAV 14

43 PHEIDISBNAV 55

RESULT 15

10 HMA\_CERVA STANDARD: PRT: 141 AA.

11 HMA\_CERVA

12 P01928:

13 01-APR-1988 (Ref: 07, Created)

14 01-APR-1988 (Ref: 07, Last sequence update)

15 01-JUN-1994 (Ref: 29, Last annotation update)

16 HMA\_CERVA ALPHA CHAIN.

17 Cebus capuchinus (white faced capuchin).

18 Eukaryotic Metazoan Chordata: Vertebrate: Eucaryotom:

19 Mammalia: Eutheria: Primates: Platyrrhini: Cebidae: Cebus.

20 NCRL\_T0010 95157

21

22 SPECIES

23 MEDLINE 86187258 PubMed 8644444

24 Tanaka Y, Arita A, Matsuda G,

25 "Primary structure of adult hemoglobin of white-throated capuchin.

26 Cebus capuchinus."

27 J Biol Chem 263:1471-1474 (1988)

28 PIR: A02294: HMA\_CERVA

29 HSSP: P01922: PHEIDISBNAV

30 InterPro: IP000971:

31 InterPro: IP000971:

32 PIR: A02294: HMA\_CERVA

33 PIR: A02294: HMA\_CERVA

34 PIR: A02294: HMA\_CERVA

35 HMA\_CERVA ALPHA CHAIN.

36 HMA\_CERVA ALPHA CHAIN.

37 HMA\_CERVA ALPHA CHAIN.

38 HMA\_CERVA ALPHA CHAIN.

39 HMA\_CERVA ALPHA CHAIN.

40 HMA\_CERVA ALPHA CHAIN.

41 HMA\_CERVA ALPHA CHAIN.

42 HMA\_CERVA ALPHA CHAIN.

43 HMA\_CERVA ALPHA CHAIN.

44 HMA\_CERVA ALPHA CHAIN.

45 HMA\_CERVA ALPHA CHAIN.

46 HMA\_CERVA ALPHA CHAIN.

47 HMA\_CERVA ALPHA CHAIN.

48 HMA\_CERVA ALPHA CHAIN.

49 HMA\_CERVA ALPHA CHAIN.

50 HMA\_CERVA ALPHA CHAIN.

Query Match 100.0%; Score 72; DB 1; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FPHEDLSH;SAQV 13  
 |||||  
 DB 43 FPHEDLSH;SAQV 55

Search completed: June 21, 2001, 10:29:19  
 Job time: 248 sec

•  
•  
•

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 21, 2001, 10:24:48 ; Search time 30.13 seconds  
(without alignments)  
30.181 Million cell updates/sec

```
Title: MS-08-832-44B-2
Perfect score: 90
Sequence: 1 CFPFDLSHGSAQVC 15
```

Scoring table: `blissum62`  
`capop 10.0` , `capext 0.5`

Searched: 412676 seqs, 60623988 residues  
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length:	0
Maximum DB seq length:	2000000000

Post processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database

1:	AI980587.qcdata/qceneseq/qceneseq/AA1980.DAT	**
2:	AI980587.qcdata/qceneseq/qceneseq/AA1981.DAT	**
3:	AI980587.qcdata/qceneseq/qceneseq/AA1982.DAT	**
4:	AI980587.qcdata/qceneseq/qceneseq/AA1983.DAT	**
5:	AI980587.qcdata/qceneseq/qceneseq/AA1984.DAT	**
6:	AI980587.qcdata/qceneseq/qceneseq/AA1985.DAT	**
7:	AI980587.qcdata/qceneseq/qceneseq/AA1986.DAT	**
8:	AI980587.qcdata/qceneseq/qceneseq/AA1987.DAT	**
9:	AI980587.qcdata/qceneseq/qceneseq/AA1988.DAT	**
10:	AI980587.qcdata/qceneseq/qceneseq/AA1989.DAT	**
11:	AI980587.qcdata/qceneseq/qceneseq/AA1990.DAT	**
12:	AI980587.qcdata/qceneseq/qceneseq/AA1991.DAT	**
13:	AI980587.qcdata/qceneseq/qceneseq/AA1992.DAT	**
14:	AI980587.qcdata/qceneseq/qceneseq/AA1993.DAT	**
15:	AI980587.qcdata/qceneseq/qceneseq/AA1994.DAT	**
16:	AI980587.qcdata/qceneseq/qceneseq/AA1995.DAT	**
17:	AI980587.qcdata/qceneseq/qceneseq/AA1996.DAT	**
18:	AI980587.qcdata/qceneseq/qceneseq/AA1997.DAT	**
19:	AI980587.qcdata/qceneseq/qceneseq/AA1998.DAT	**
20:	AI980587.qcdata/qceneseq/qceneseq/AA1999.DAT	**
21:	AI980587.qcdata/qceneseq/qceneseq/AA2000.DAT	**
22:	AI980587.qcdata/qceneseq/qceneseq/AA2001.DAT	**

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARY

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	50	100.0	15	18	AAW30116	Cytosine peptide 433
	2	75.5	88.3	16	17	AAK67246	Peptide fragment c
	3	72	80.0	13	18	AAW30336	Peptide 43-55 with
	4	72	80.0	16	20	AAV41983	Rheumatoid arthritis
	5	72	80.0	19	19	AAW66276	Circulantly formidate
	6	72	80.0	14	13	AAK29606	Mutant human alpha
	7	72	80.0	14	13	AAK29608	Mutant human alpha
	8	72	80.0	14	14	AAK30533	Titusville alpha c
	9	72	80.0	14	14	AAK42283	Human alpha Haemoco
	10	72	80.0	14	14	AAK42681	Neutral alpha-glob
	11	72	80.0	14	15	AAK55131	Human alpha haemoco

12	72	80.0	141	15	AAK44516	Adult human alpha-1
13	72	80.0	141	15	AAK44517	Adult human alpha-1
14	72	80.0	141	15	AAK44518	Adult human alpha-1
15	72	80.0	141	15	AAK44519	Adult human alpha-1
16	72	80.0	141	15	AAK44520	Adult human alpha-1
17	72	80.0	141	17	AAK97359	Human alpha haemoglobin alpha chain of
18	72	80.0	141	18	AAK40312	The alpha chain of
19	72	80.0	141	18	AAK40734	Haemoglobin alpha
20	72	80.0	141	18	AAK40734	Haemoglobin alpha
21	72	80.0	141	18	AAK40731	Native human haemo-
22	72	80.0	141	18	AAK40724	Haemoglobin alpha
23	72	80.0	141	18	AAK40726	Haemoglobin alpha
24	72	80.0	141	18	AAK40726	Haemoglobin alpha
25	72	80.0	141	18	AAK40716	Recombinant wild t
26	72	80.0	141	18	AAK10470	Human alpha-haemo-
27	72	80.0	141	18	AAK11868	Human alpha-haemo-
28	72	80.0	141	18	AAK73389	Human haemoglobin
29	72	80.0	141	20	AAK73386	Human haemoglobin
30	72	80.0	141	20	AAK94056	Alpha chain of human
31	72	80.0	141	20	AAK13376	Human alpha-haemo-
32	72	80.0	141	20	AAK44541	Human alpha-haemo-
33	72	80.0	141	21	AAK87794	Human alpha-haemo-
34	72	80.0	141	21	AAK79379	Human alpha-haemo-
35	72	80.0	141	22	AAK64774	Human mutant alpha
36	72	80.0	142	20	AAK94058	Human haemoglobin a
37	72	80.0	283	14	AAK39457	Human haemoglobin (
38	72	80.0	282	19	AAK54282	Protein sequence of
39	72	80.0	284	11	AAK08114	(Glu-Val)-alpha (C
40	72	80.0	285	19	AAK54283	75th sequence of the
41	72	80.0	573	19	AAK41284	DNA sequence of the
42	71	78.9	571	14	AAK45219	Human adult alpha
43	71	78.9	571	14	AAK45219	IIIa, Mr. 42,000, haemo-
44	71	78.9	571	14	AAK45219	globin, oxygen affinity
45	71	78.9	141	16	AAK70979	Crocodile alpha-1

## ALIGNMENTS

	RESULT	1
AAW0336	standard; peptide; 15 AA.	
ID	AAW03016:	
XX		
AC		
XX		
DT	20-APR-1998 (first entry)	
XX		
XX	Cellular factor 18.57 which inhibits/proliferative activity.	
FW	INPRED stem cells inhibiting stimulation proliferation myoblasts	
FW	alpha-defensin associated with cell death and apoptosis	
KM	differential protection stem cell hypoproliferation aplastic anaemia	
AM	stem cell exhaustion cyclic peptide.	
XX		
OS	Synthetic.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FF	Disulfide bond 1..15	
XX		
FN	W09736922-A1.	
XX		
PID	09-OCT-1997.	
AX		
PP	03-APR-1997: 97MO-US05601.	
XX		
FE	03 APR-1996: 96US 0627173.	
XX		
PA	(PRON ) PRO NEURON INC.	
XX		
PI	Tsylova I., Wolpe SD;	
XX		
DR	WP1 : 1997-503044/416.	



DR WP1: 1997-50044/46.

XX PS Haemoglobin alpha chain derivatives - used to inhibit or stimulate  
PT stem cell proliferation, for treatment of cancer, immune depression  
PT etc  
XX PS Claim 46: Page 116: 162pp; English.

XX CC The present peptide is an INP60L peptide that has stimulatory and  
CC inhibitory activity on stem cell proliferation. The 3-dimensional  
CC structure of myoglobin was superimposed onto the 3-dimensional  
CC structure of the alpha chain present in human haemoglobin, and the  
CC present peptide identified as a region which is structurally different  
CC from myoglobin in 3-dimensional space. The peptide is used to  
CC inhibit stem cell proliferation, particularly during radiotherapy or  
CC chemotherapy of cancer, but more generally wherever a mammal is exposed  
CC to an agent, e.g. an antiviral, that damages or destroys stem cells. It  
CC can also be used for antiviral protection of normal stem cells, but  
CC not cancer cells, from chemotherapy or radiation, particularly after  
CC normal stem cells have been induced to proliferate by therapeutic  
CC treatment. The peptide can also be used for control of stem cell  
CC hypoproliferation (e.g. aplastic anaemia), for treating or preventing  
CC stem cell exhaustion (e.g. where caused by acquired immune deficiency  
CC syndrome), and to treat or prevent immunodeficiency. INP60L can also be  
CC used for treatment of a mammal, adult and juvenile compounds, so  
CC reversibly inhibit or stimulate stem cells depending on the dose, so  
CC allow precise control over cycling of these cells.

XX S0 Sequence 13 AA:

Query Match H0.0% Score 72; DB 18; Length 13;

Best Local Similarity: 100.0%; Pred. No. 2.2e-06;

Matches: 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 Iphidshasay 13

QY 2 FPHHDLHSNAYV 14

|||||

DB 1 Iphidshasay 13

RESULT 4

AAV41983 standard; Peptide: 16 AA.

XX AC AAV41983;

XX DT 09-DEC-1999 (first entry)

XX DE Rheumatoid arthritis diagnostic protein isoform peptide #14

XX KM human; rheumatoid arthritis; RA; diagnosis; RPI; RAPI; detection;

XX KM rheumatoid arthritis diagnostic feature; RPI; synovial fluid;

XX KM rheumatoid arthritis diagnostic protein isoform; sequencing;

XX OS Homo sap. ens.

XX PN W09947925-A2.

XX PD 23-SEP-1999.

XX PF 15-MAR-1999; 94W0-GB00763.

XX PR 13-MAR-1998; 98GB-0005477.

XX PA (OXFO-) XFOOD GLYC/SCIENCES UK LTD.

XX PT Parekh RB, Patel TP, Townsend RR;

XX WP1: 1999-571871/48

XX PT diagnosis of human rheumatoid arthritis by two-dimensional  
PT electrophoresis -

XX PS Disclosure: Page 20; 157pp; English.

XX CC A method has been developed for the diagnosis of human rheumatoid  
CC arthritis (RA) using two-dimensional electrophoresis to generate a  
CC two dimensional array of features. The method can be used for screening  
CC diagnosis and prognosis of RA in a subject or for monitoring the effect  
CC of an anti-RA drug or therapy administered to a subject. The method  
CC comprises: (a) analysing a sample of serum or plasma and optionally  
CC synovial fluid by two dimensional electrophoresis, to generate a two-  
CC dimensional array of features; (b) identifying at least one chosen  
CC feature whose relative abundance correlates with the presence or absence  
CC of RA; and (c) comparing the abundance of each chosen feature in the  
CC sample with the abundance of that chosen feature in serum or plasma from  
CC one or more persons without RA, where the relative abundance of the  
CC chosen feature or features in the sample indicates the presence or  
CC absence of RA in the subject. The method can also be used in clinical  
CC studies for testing drugs for therapy of RA, for purification of RA-  
CC diagnostic protein isoforms (RPIs), and for production of antibodies to  
CC RPIs. The RA diagnostic feature (RAPI) proteins can be used to identify  
CC compounds that promote or inhibit their activity, which are then used as  
CC RA drugs. Nucleic acid encoding RAPIs can be used in gene therapy.  
CC Proteolysis: AAV41944 to AAV42103 represent RPI peptides; AAV42101 to  
CC AAV42103 represent expression reference protein isoform peptides and  
CC AAV42008 to AAV42092 represent degenerate protein RPIs, which are all  
CC used in the exemplification of the present invention.

XX S0 Sequence 16 AA:

Query Match H0.0% Score 72; DB 20; Length 16;

Best Local Similarity: 100.0%; Pred. No. 2.7e-06;

Matches: 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 3 Iphidshasay 16

QY 2 FPHHDLHSNAYV 14

|||||

DB 3 Iphidshasay 16

RESULT 5

AAW66275 standard; peptide: 19 AA.

XX AC AAW66275;

XX DT 23-NOV-1998 (first entry)

XX DE circularly permuted  $\alpha$ -alpha haemoglobin-like peptide fragment.

XX KM haemoglobin; circularly permuted; N terminus; C terminus; multi-letter;

XX KM permuted haemoglobin; globin; PDB; red blood cells; blood transfusion;

XX OS Synthetic.

XX PN W09838211-A2.

XX PD 03-SEP-1998.

XX PF 27-FEB-1998; 98W0-050836.

XX PR 28-FEB-1997; 97DS-0039657.

XX FA (SOMA-) SOMATOSTATIN 1N.

XX PT Best EA, cllins PO;

XX WP1: 1998-481141/41.

XX PT New haemoglobin form with non-naturally occurring termini  
PT useful for the treatment of any condition where therapy involves red  
PT blood cells or oxygen delivery









PN	WC9325071-A.
XX	
PD	23-DEC-1993.
XX	
PF	11-JUN-1993; 94WO-0805629.
XX	
PR	12-JUN-1992; 92US-0897648.
PR	08-MAR-1992; 92US-0987890.
PR	15-MAR-1993; 94US-0030897.
XX	
PA	(ONXU-) INX (CKRP).
XX	
P1	Holtzman SH, Koller H, Kumar R, Lofan JS, O' Donnell JK;
P1	Parsons CT, Pilder SH, Plinkert CA, Sharma A, Swanson MB;
P1	White ST.
XX	
DR	WPI: 1994-007075/01.
XX	
PT	Transgenic pigs which produce human haemoglobin - cont'd. DNA encoding
PT	human alpha globin and huma beta globin linked to promoters
XX	
PS	Claim 14 : 117pp; English.
XX	
C1	The sequences given in AAR44517-20 represent mutant versions of the
C1	adult human alpha globin sequence. These sequences were not given in
C1	the specification but were generated using the known alpha globin
C1	sequence. These sequences were used in the production of constructs which
C1	were used to produce the transgenic pigs of the invention. These
C1	constructs contain the pig beta globin promoter sequence operably linked
C1	to the human alpha or beta globin genes. Pigs containing these
C1	constructs express human hemoglobin (Hb) in their erythrocytes and are
C1	healthy, suffering no deleterious side effects as a result of
C1	heterologous Hb production. They can be used as an efficient and
C1	economical source of human Hb that can be used for transfusion and other
C1	medical applications. Also, a human-pig hybrid Hb can be produced which
C1	exhibits a p50 that is higher than that of native human or pig Hb.
SO	Sequence 141 AA:
OY	2 FPHDLSHGSAV 14       Matches 14; conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	43 tptdlshtsqav 55
RESULT 13	
AAR44517	Local Seq./Consiliones
ID	AAR44517 standard; protein; 141 AA.
XX	
AC	AAR44517;
XX	
DP	23-JUN-1994 (first entry)
DE	
DE	Adult human alpha-globin E36Y.
KW	Adult; pi(g); beta; globin gene; regulatory region; promoter;
KW	transcn c pig; human, alpha, haemoglobin; Hb; erythrocyte;
KW	transcn c; human; T53; h_1 r13 nt.
OS	homo sapiens.
XX	
PII	key
PII	Misc-difference 36
PII	/label= E36y
PII	WG9325071-A.
PD	23-DEC-1993
XX	

[illegible]





[illegible]

Quartz	Mass %	100.00	Trace	7.7	100.00	Length	141
Host	Local stability	100.00	Trace	None	100.00		
Mutual	Mass %	0.00	Mutual	0.00	0.00	0.00	0.00
Q2	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
Q4	4	4	4	4	4	4	4

RESIDUE	SEQUENCE	PROTEIN NAME	PROT	AA
1	QVYRZ	PROTEIN	142	AA
2	QVYRZ	PROTEIN	142	AA
3	QVYRZ	PROTEIN	142	AA
4	QVYRZ	PROTEIN	142	AA
5	QVYRZ	PROTEIN	142	AA
6	QVYRZ	PROTEIN	142	AA
7	QVYRZ	PROTEIN	142	AA
8	QVYRZ	PROTEIN	142	AA
9	QVYRZ	PROTEIN	142	AA
10	QVYRZ	PROTEIN	142	AA
11	QVYRZ	PROTEIN	142	AA
12	QVYRZ	PROTEIN	142	AA
13	QVYRZ	PROTEIN	142	AA
14	QVYRZ	PROTEIN	142	AA
15	QVYRZ	PROTEIN	142	AA
16	QVYRZ	PROTEIN	142	AA
17	QVYRZ	PROTEIN	142	AA
18	QVYRZ	PROTEIN	142	AA
19	QVYRZ	PROTEIN	142	AA
20	QVYRZ	PROTEIN	142	AA
21	QVYRZ	PROTEIN	142	AA
22	QVYRZ	PROTEIN	142	AA
23	QVYRZ	PROTEIN	142	AA
24	QVYRZ	PROTEIN	142	AA
25	QVYRZ	PROTEIN	142	AA
26	QVYRZ	PROTEIN	142	AA
27	QVYRZ	PROTEIN	142	AA
28	QVYRZ	PROTEIN	142	AA
29	QVYRZ	PROTEIN	142	AA
30	QVYRZ	PROTEIN	142	AA
31	QVYRZ	PROTEIN	142	AA
32	QVYRZ	PROTEIN	142	AA
33	QVYRZ	PROTEIN	142	AA
34	QVYRZ	PROTEIN	142	AA
35	QVYRZ	PROTEIN	142	AA
36	QVYRZ	PROTEIN	142	AA
37	QVYRZ	PROTEIN	142	AA
38	QVYRZ	PROTEIN	142	AA
39	QVYRZ	PROTEIN	142	AA
40	QVYRZ	PROTEIN	142	AA
41	QVYRZ	PROTEIN	142	AA
42	QVYRZ	PROTEIN	142	AA
43	QVYRZ	PROTEIN	142	AA
44	QVYRZ	PROTEIN	142	AA
45	QVYRZ	PROTEIN	142	AA
46	QVYRZ	PROTEIN	142	AA
47	QVYRZ	PROTEIN	142	AA
48	QVYRZ	PROTEIN	142	AA
49	QVYRZ	PROTEIN	142	AA
50	QVYRZ	PROTEIN	142	AA
51	QVYRZ	PROTEIN	142	AA
52	QVYRZ	PROTEIN	142	AA
53	QVYRZ	PROTEIN	142	AA
54	QVYRZ	PROTEIN	142	AA
55	QVYRZ	PROTEIN	142	AA
56	QVYRZ	PROTEIN	142	AA
57	QVYRZ	PROTEIN	142	AA
58	QVYRZ	PROTEIN	142	AA
59	QVYRZ	PROTEIN	142	AA
60	QVYRZ	PROTEIN	142	AA
61	QVYRZ	PROTEIN	142	AA
62	QVYRZ	PROTEIN	142	AA
63	QVYRZ	PROTEIN	142	AA
64	QVYRZ	PROTEIN	142	AA
65	QVYRZ	PROTEIN	142	AA
66	QVYRZ	PROTEIN	142	AA
67	QVYRZ	PROTEIN	142	AA
68	QVYRZ	PROTEIN	142	AA
69	QVYRZ	PROTEIN	142	AA
70	QVYRZ	PROTEIN	142	AA
71	QVYRZ	PROTEIN	142	AA
72	QVYRZ	PROTEIN	142	AA
73	QVYRZ	PROTEIN	142	AA
74	QVYRZ	PROTEIN	142	AA
75	QVYRZ	PROTEIN	142	AA
76	QVYRZ	PROTEIN	142	AA
77	QVYRZ	PROTEIN	142	AA
78	QVYRZ	PROTEIN	142	AA
79	QVYRZ	PROTEIN	142	AA
80				

Country	Method	Year	Sample Size	Length
United States	Longitudinal	1972-1973	1,000	14.2
United States	Longitudinal	1974-1975	1,000	14.2
United States	Longitudinal	1976-1977	1,000	14.2
United States	Longitudinal	1978-1979	1,000	14.2
United States	Longitudinal	1980-1981	1,000	14.2
United States	Longitudinal	1982-1983	1,000	14.2
United States	Longitudinal	1984-1985	1,000	14.2
United States	Longitudinal	1986-1987	1,000	14.2
United States	Longitudinal	1988-1989	1,000	14.2
United States	Longitudinal	1990-1991	1,000	14.2
United States	Longitudinal	1992-1993	1,000	14.2
United States	Longitudinal	1994-1995	1,000	14.2
United States	Longitudinal	1996-1997	1,000	14.2
United States	Longitudinal	1998-1999	1,000	14.2
United States	Longitudinal	2000-2001	1,000	14.2
United States	Longitudinal	2002-2003	1,000	14.2
United States	Longitudinal	2004-2005	1,000	14.2
United States	Longitudinal	2006-2007	1,000	14.2
United States	Longitudinal	2008-2009	1,000	14.2
United States	Longitudinal	2010-2011	1,000	14.2
United States	Longitudinal	2012-2013	1,000	14.2
United States	Longitudinal	2014-2015	1,000	14.2
United States	Longitudinal	2016-2017	1,000	14.2
United States	Longitudinal	2018-2019	1,000	14.2
United States	Longitudinal	2020-2021	1,000	14.2
United States	Longitudinal	2022-2023	1,000	14.2
United States	Longitudinal	2024-2025	1,000	14.2
United States	Longitudinal	2026-2027	1,000	14.2
United States	Longitudinal	2028-2029	1,000	14.2
United States	Longitudinal	2030-2031	1,000	14.2
United States	Longitudinal	2032-2033	1,000	14.2
United States	Longitudinal	2034-2035	1,000	14.2
United States	Longitudinal	2036-2037	1,000	14.2
United States	Longitudinal	2038-2039	1,000	14.2
United States	Longitudinal	2040-2041	1,000	14.2
United States	Longitudinal	2042-2043	1,000	14.2
United States	Longitudinal	2044-2045	1,000	14.2
United States	Longitudinal	2046-2047	1,000	14.2
United States	Longitudinal	2048-2049	1,000	14.2
United States	Longitudinal	2050-2051	1,000	14.2
United States	Longitudinal	2052-2053	1,000	14.2
United States	Longitudinal	2054-2055	1,000	14.2
United States	Longitudinal	2056-2057	1,000	14.2
United States	Longitudinal	2058-2059	1,000	14.2
United States	Longitudinal	2060-2061	1,000	14.2
United States	Longitudinal	2062-2063	1,000	14.2
United States	Longitudinal	2064-2065	1,000	14.2
United States	Longitudinal	2066-2067	1,000	14.2
United States	Longitudinal	2068-2069	1,000	14.2
United States	Longitudinal	2070-2071	1,000	14.2
United States	Longitudinal	2072-2073	1,000	14.2
United States	Longitudinal	2074-2075	1,000	14.2
United States	Longitudinal	2076-2077	1,000	14.2
United States	Longitudinal	2078-2079	1,000	14.2
United States	Longitudinal	2080-2081	1,000	14.2
United States	Longitudinal	2082-2083	1,000	14.2
United States	Longitudinal	2084-2085	1,000	14.2
United States	Longitudinal	2086-2087	1,000	14.2
United States	Longitudinal	2088-2089	1,000	14.2
United States	Longitudinal	2090-2091	1,000	14.2
United States	Longitudinal	2092-2093	1,000	14.2
United States	Longitudinal	2094-2095	1,000	14.2
United States	Longitudinal	2096-2097	1,000	14.2
United States	Longitudinal	2098-2099	1,000	14.2
United States	Longitudinal	2100-2101	1,000	14.2
United States	Longitudinal	2102-2103	1,000	14.2
United States	Longitudinal	2104-2105	1,000	14.2
United States	Longitudinal</			

```

1  PPPHSHSSAY 14
2  | | | | | | | |
3  44 PPPHSHSSAY 50

```

110	Q2874	PRELIMINARY:	Feb: 14. AA.
111	Q2874		
112	01 NOV 1996	(TEMBREL, 01, closed)	
113	01 NOV 1996	(TEMBREL, 01, last sequence update)	
114	01 MAR 2001	(TEMBREL, 16, last annotation update)	
115	ALPHA-GLOBIN CHAIN:		
116	OVERLAP (Sheep):		
117	EXON/INTRON: Metazoa:		
118	EXON/INTRON: Metazoa:		
119	EXON/INTRON: Metazoa:		
120	EXON/INTRON: Metazoa:		
121	EXON/INTRON: Metazoa:		
122	EXON/INTRON: Metazoa:		
123	EXON/INTRON: Metazoa:		
124	EXON/INTRON: Metazoa:		
125	EXON/INTRON: Metazoa:		
126	EXON/INTRON: Metazoa:		
127	EXON/INTRON: Metazoa:		
128	EXON/INTRON: Metazoa:		
129	EXON/INTRON: Metazoa:		
130	EXON/INTRON: Metazoa:		
131	EXON/INTRON: Metazoa:		
132	EXON/INTRON: Metazoa:		
133	EXON/INTRON: Metazoa:		
134	EXON/INTRON: Metazoa:		
135	EXON/INTRON: Metazoa:		
136	EXON/INTRON: Metazoa:		
137	EXON/INTRON: Metazoa:		
138	EXON/INTRON: Metazoa:		
139	EXON/INTRON: Metazoa:		
140	EXON/INTRON: Metazoa:		
141	EXON/INTRON: Metazoa:		
142	EXON/INTRON: Metazoa:		
143	EXON/INTRON: Metazoa:		
144	EXON/INTRON: Metazoa:		
145	EXON/INTRON: Metazoa:		
146	EXON/INTRON: Metazoa:		
147	EXON/INTRON: Metazoa:		
148	EXON/INTRON: Metazoa:		
149	EXON/INTRON: Metazoa:		
150	EXON/INTRON: Metazoa:		

	100.0%	Score 7.2	148 of 149	Length 142:
	Post Local Similarity:	100.0%,	Prod. No. 6, "no-06"	
Motif-class	12	"conservative"	01	Motif-classes 01
				Indels 01
Cy	1	PREFDISHSNAGV 14		
Ld	44	PREFDISHSNAGV 56		
PSMID	5			
G2H744				
D-	G2H744	PRELIMINARY:	Pkt: 142 AA.	
A*	G2H744:			
I 1	01 NOV 1996 (ITEMBROL_01, created)			
I 1	01 NOV 1996 (ITEMBROL_01, last sequence update)			
I 1	01 MAR 2001 (ITEMBROL_16, last amount from update)			
LE	ALPHA GLOBIN CHAIN,			
-CS	ovis aries (Sheep):			
oc	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
oc	Mammalia; Perborlia; Cetartiodactyla; Kunitaria; Bovidae; Bovinae;			
oc	Bovidae; Caprinae; Ovis:			
OX	NRL_TAXID 9940:			
RN	[1]			
KO	SOURCE: FROM N.A.			
BO	JUSSE, 1846-09:			
PX	MEDLINE 9528398: pubmed-7769611:			
R1	Kistaldi M., S.; Casula S.; Rando A.; Vestri R.:			
R1	"Sheep alpha-globin gene sequences: implications for their conserved			
R1	evolution and for the down regulation of the 5' genes.";			
R1	J Mol. Evol. 40:449-454(1995).			
CV	1 SIMILARITY: TO GLOBIN FAMILY			
CV	1 SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.			
FM02	XZGZLVZ *AAAY9760.1:			

















```

1  STATE: CAI
2  ZIP: 94111
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: Floppy disk
5  OPERATING SYSTEM: pc compat dlo
6  SOFTWARE: Format 5 Release #1.0, Version #1.05
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/92/792,692
9  FILING DATE: 11 JUN 1992
10 CLASSIFICATION: 45
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 600,244
13 FILING DATE: 22 OCT 1990
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: US 611,877
16 FILING DATE: 15 JAN 1991
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 610,981
19 FILING DATE: 17 FEB 1989
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US 600,766
22 FILING DATE: 25 FEB 1988
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 604,771
25 FILING DATE: 26 FEB 1988
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 617,637
28 FILING DATE: 05 MAY 1984
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 604,146
31 FILING DATE: 08 JUN 1989
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: US 619,274
34 FILING DATE: 21 JUL 1988
35 ATTORNEY/AGENT INFORMATION:
36 NAME: Bulfinch, Albert P.
37 REPRESENTATION NUMBER: 28,967
38 REFERENCE/POWER KEY NUMBER: B106, 201, 21
39 REFERENCE/POWER KEY NUMBER: USA
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 415, 444, 4150
42 TELEFAX: 415, 444, 8716
43 INFORMATION FOR SEQ ID NO: 1:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 141 amino acids
46 TYPE: amino acid
47 Topology: linear
48 MOLECULE TYPE: Protein
49 US 07 920 0950 1

```

Query Match: 80.0% Score 72. DB 1: Length 141  
 Best Local Similarity: 100.0% Prod. No. 832-05  
 Matches: 13 Conserved 02 Mismatches 02 Indels 02 Gaps 02

```

59  2 PPTPUSHSAV 14
60  1111111111
61  44 PPTPUSHSAV 55

```

RESULT 8  
 US 08 170 0950 1  
 Sequence 1: Application US/081700958  
 Patent No. 556,254  
 GENERAL INFORMATION:  
 APPLICANT: Bulfinch, Stephen J.  
 APPLICANT: Nagata, Kiyoshi  
 TITLE OF INVENTION: Blood Substitution  
 NUMBER OF SEQUENCES: 66  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Bulfinch, Stephen J.  
 STREET: 2001 Central Avenue

```

1  CITY: Boulder
2  STATE: Colorado
3  ZIP: 80501
4  COMPUTER READABLE FORM:
5  MEDIUM TYPE: Diskette, 5.25 inch, 1.4 MB storage
6  OPERATING SYSTEM: Apple Macintosh
7  SOFTWARE: Microsoft Word 5.04
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER: US/92/792,692
10 FILING DATE: December 20, 1994
11 CLASSIFICATION: 540
12 ATTORNEY/AGENT INFORMATION:
13 NAME: No. 55632540K, Henry P.
14 REGISTRATION NUMBER: 83200
15 REFERENCE/POWER KEY NUMBER: B11, 201, 202
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 303, 441, 3322
18 TELEFAX: 303, 444, 3014
19 INFORMATION FOR SEQ ID NO: 1:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 141
22 TYPE: amino acid
23 STRANDNESS: single
24 Topology: linear
25 MOLECULE TYPE: Protein
26 US 08 170 0950 1

```

Query Match: 80.0% Score 72. DB 1: Length 141  
 Best Local Similarity: 100.0% Prod. No. 832-05  
 Matches: 13 Conserved 02 Mismatches 02 Indels 02 Gaps 02

```

59  2 PPTPUSHSAV 14
60  1111111111
61  44 PPTPUSHSAV 55

```

RESULT 9  
 US 08 184 227 6  
 Sequence 8: Application US/08184227  
 Patent No. 5589,667  
 GENERAL INFORMATION:  
 APPLICANT: Hanson, Jon  
 APPLICANT: Lawson, William O.  
 APPLICANT: Graham, George L.  
 APPLICANT: Turpen, Thomas H.  
 APPLICANT: Garney, Ann Myers  
 APPLICANT: Garney, Stephen J.  
 TITLE OF INVENTION: REGENERANT PLANT VITAL NUTRIENT ACTORS  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Limbach & Limbach  
 STREET: 2001 Betty Building  
 CITY: San Francisco  
 STATE: CAI  
 ZIP: 94111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: pc compat dlo  
 SOFTWARE: Format 5 Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/98/184,227  
 FILING DATE:  
 CLASSIFICATION: 45  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 925,692  
 FILING DATE: 01 JUL 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 600,244  
 FILING DATE: 22 OCT 1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 641,617  
FILING DATE: 16-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 310,891  
FILING DATE: 17-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,766  
FILING DATE: 26-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,771  
FILING DATE: 28-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 347,637  
FILING DATE: 05-MAY-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 463,138  
FILING DATE: 08-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 219,279  
FILING DATE: 15-JUL-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Hallibur, Albert P.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 4103-20121 USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-433-4150  
TELEFAX: 415-433-8716  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-184-237-8

Query Match: 80.0%, Score 72, PB 1, Length 141;  
Host Local Similarity: 100.0%, Pred. No. 5,36-05;

Matches: 13, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0;

QY 2 FPHDLSHSAV 14  
ID 43 FPHDLSHSAV 55

RESULT 10  
US-08-240-712-17  
Sequence 17, Application US/08-240712  
Patent No. 5599907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID G.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STEHLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Broadway and Belmont  
STREET: 419 Seventh Street, N.W., Suite 400  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/99-240-712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 510

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: AMALPSOR-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-727-9528  
TELEX: 248643  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-240-712-17

Query Match: 80.0%, Score 72, PB 1, Length 141;

Host Local Similarity: 100.0%, Pred. No. 5,36-05;

Matches: 13, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0;

QY 2 FPHDLSHSAV 14  
ID 43 FPHDLSHSAV 55

RESULT 11  
US-08-240-712-26  
Sequence 26, Application US/08-240712  
Patent No. 5599907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID G.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STEHLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Broadway and Belmont  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/98-240-712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 510  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: AMALPSOR-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-727-9528  
TELEX: 248643  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single





SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/443,890  
APPLICATION NUMBER: US/08/443,890  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 510  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,712  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: 1-7-70542,700752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-443-890-17

Query Match 80.0% Score 72; DB 1; Length 141;  
Best local similarity 100.0%; Pred. No. 5,46-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHFDUSHSAGV 14  
|||||  
DB 43 FPHFDUSHSAGV 55

RESULT 15  
US-08-443-890-26  
Sequence 26, Application US/08443890  
Patent No. 5739011  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STELLER, GARY L.  
TITLE OF INVENTION: PREVENTION AND USE OF MULTIMERIC  
TITLE OF INVENTION: HEMOGLOBINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESS: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/443,890  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 510  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,712  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: 1-7-70542,700752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON-6

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-443-890-26

Query Match 80.0% Score 72; DB 1; Length 141;  
Best local similarity 100.0%; Pred. No. 5,46-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHFDUSHSAGV 14  
|||||  
DB 43 FPHFDUSHSAGV 55

Search completed: June 21, 2001, 10:25:09  
Job time: 258 sec

•

•

•





R:Mahoney, W.C.; Note, P.H.  
Arch. Biochem. Biophys. 196, 64-72, 1979  
A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(1)  
A:Reference number: S06514; MUID:80063760  
A:Accession: S06514  
A:Molecule type: protein  
A:Residues: 1-141<MAT>  
C:Superfamily: globin: globin homology  
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterodimer; iron; metalloprotein  
F:2-141/Domain: globin homology <GLB>  
F:58/Binding site: oxygen (His) (distal axial ligand) \*status predicted  
F:87/Binding site: heme iron (His) (proximal axial ligand) \*status predicted

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

RESULT 8  
HAMQ2P  
hemoglobin alpha-1 chain: pig-tailed macaque  
C:Species: Macaca nemestrina (pig-tailed macaque)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 03-Mar-2000  
A:Accession: S07670  
R:Mahoney, W.C.; Note, P.H.  
Arch. Biochem. Biophys. 196, 64-72, 1979  
A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(1)  
A:Reference number: S06514; MUID:80063760  
A:Accession: S07670  
A:Molecule type: protein  
A:Residues: 1-141<MAH>  
C:Superfamily: globin: globin homology  
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterodimer; iron; metalloprotein  
F:2-141/Domain: globin homology <GLB>  
F:58/Binding site: oxygen (His) (distal axial ligand) \*status predicted  
F:87/Binding site: heme iron (His) (proximal axial ligand) \*status predicted

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

RESULT 9  
HAMKP  
hemoglobin alpha chain: black handed spider monkey  
C:Species: Atelax geoffroyi (black-handed spider monkey)  
C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 03-Mar-2000  
A:Accession: A02253  
R:Matsuda, G.; Matita, T.; Sugiyama, Y.; Setoyoshi, M.; Ota, Y.; Araya, A.; Gohtani, M.;  
Hoppe-Seyler, K.; Physiol. Chem. 354, 1573-1576, 1973  
A:Title: Studies on the primary structures of alpha and beta polypeptide chains of adult  
A:Reference number: A02253; MUID:75059125  
A:Accession: A02253  
A:Molecule type: protein  
A:Residues: 1-141<MAT>  
C:Superfamily: globin: globin homology  
C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier  
F:2-141/Domain: globin homology <GLB>  
F:58/Binding site: oxygen (His) (distal axial ligand) \*status predicted  
F:87/Binding site: heme iron (His) (proximal axial ligand) \*status predicted

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

RESULT 10  
HACJB  
hemoglobin alpha chain: black tailed marmoset  
C:Species: Callithrix argentata (black-tailed marmoset)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 03-Mar-2000  
A:Accession: S06512  
R:Matita, T.; Hayashida, M.; Matsuda, G.  
J. Biochem. 95, 805-813, 1984  
A:Title: Primary structures of adult hemoglobins of Callithrix argentata  
A:Reference number: S06512; MUID:84212383  
A:Accession: S06512  
A:Molecule type: protein  
A:Residues: 1-141<MAT>  
C:Superfamily: globin: globin homology  
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterodimer; iron; metalloprotein  
F:2-141/Domain: globin homology <GLB>  
F:58/Binding site: oxygen (His) (distal axial ligand) \*status predicted  
F:87/Binding site: heme iron (His) (proximal axial ligand) \*status predicted

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

RESULT 11  
HAMQA  
hemoglobin alpha chain: brown capuchin  
C:Species: Cebus apella (brown capuchin)  
C:Date: 28-May-1985 #sequence\_revision 28-May-1985 #text\_change 03-Mar-2000  
A:Accession: A02254  
R:Matsuda, G.; Matita, T.; Watanabe, R.; Araya, A.; Matsuda, K.; Ota, Y.; Goodman, M.;  
Hoppe-Seyler, K.; Physiol. Chem. 354, 1513-1516, 1973  
A:Title: The amino acid sequences of the alpha and beta polypeptide chains of adult h  
A:Reference number: A02254; MUID:75059125  
A:Accession: A02254  
A:Molecule type: protein  
A:Residues: 1-141<MAT>  
C:Superfamily: globin: globin homology  
C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen car  
F:2-141/Domain: globin homology <GLB>  
F:58/Binding site: oxygen (His) (distal axial ligand) \*status predicted  
F:87/Binding site: heme iron (His) (proximal axial ligand) \*status predicted

Query Match 80.0%; Score 72; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 PFPHDSHSAOV 14  
|||||  
43 PFPHDSHSAOV 55

RESULT 12  
HAMQF  
hemoglobin alpha chain: brown headed tamarin  
C:Species: Saguinus fuscicollis (brown-headed tamarin, saddle-backed tamarin)  
C:Date: 28-May-1985 #sequence\_revision 28-May-1985 #text\_change 03-Mar-2000  
A:Accession: A02255  
R:Lin, K.D.; Kim, Y.K.; Chernoff, A.L.

hemoglobin, Genet. 14: 427-440, 1976.

At E111: The primary structure of the haemoglobin (status: HIS) hemoglobin, E. Use of

At Reference number: A02250, M010:7021459

At Accession: A02250

At Residues: 1141 - 1141

At Superfamily: globin, alpha hemoglobin

At Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier

E12:141/Domains: globin hemoglobin - GHB

E12/Binding sites: oxygen (HIS) (distal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

Best Local Similarity: 100.0% (Prod. No. 60-05)  
Matches: 133 Conserved: 02 Mismatches: 02 Indels: 02 Gaps: 02

Query: 2 EPHB1SHSAAV 14

Id: 43 EPHB1SHSAAV 55

Result: 15

HAIRN

hemoglobin alpha chain - slow tortis

At Species: Nymphaeus concolor (slow tortis)

At Date: 24 Apr 1984 #sequence revision 41 Dec 1991 #text change 04 Mar 2000

At Accession: A02260

At Residues: 419-421, 1973

At Superfamily: globin, alpha hemoglobin

At Keywords: blood; chromoprotein; erythrocyte; heme; hemeoelctamer; iron; metallopro

E12:141/Domains: globin hemoglobin - GHB

E12/Binding sites: oxygen (HIS) (distal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

Search completed: June 21, 2001, 10:25:33

Job time: 15.2 sec

Query Match: 80.0% Score 72: 18 12 Length 141:

Best Local Similarity: 100.0% (Prod. No. 60-05)

Matches: 133 Conserved: 02 Mismatches: 02 Indels: 02 Gaps: 02

Query: 2 EPHB1SHSAAV 14

Id: 43 EPHB1SHSAAV 55

Search completed: June 21, 2001, 10:25:33

Job time: 15.2 sec

E12/Binding sites: heme iron (HIS) (proximal axial ligand) \*status predicted

GenCore version 4.5  
Copyright (c) 1994 - 2000 Compugen Ltd.

OM protein protein search, using sw model

Run on: June 21, 2001, 10:29:19 : Search time 11.96 Seconds

(without alignments)  
42.663 Million col1 updates/sec

File: US-08-832-443b-2

Perfect score: 50

Sequence: 1 FPHFDLSHSAQVC 15

Scoring table:  
BLSDM62  
Gapop 10 0, Gapext 0 5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	80.0	140	HBA_TRAFT	P04237 Trachelaphus
2	72	80.0	141	HBA_TRAFT	P1755C Trachelaphus
3	72	80.0	141	HBA_TRAFT	P1330 Trachelaphus
4	72	80.0	141	HBA_TRAFT	P03668 bos mutus s
5	72	80.0	141	HBA_TRAFT	P01921 alces alces
6	72	80.0	141	HBA_TRAFT	P14387 antilocapra p
7	72	80.0	141	HBA_TRAFT	P01927 alces alces
8	72	80.0	141	HBA_TRAFT	P09423 bison bonas
9	72	80.0	141	HBA_TRAFT	P01969 bos taurus
10	72	80.0	141	HBA_TRAFT	P01966 bos taurus
11	72	80.0	141	HBA_TRAFT	P18973 capra hircu
12	72	80.0	141	HBA_TRAFT	P01970 capra hircu
13	72	80.0	141	HBA_TRAFT	P01928 cebus apell
14	72	80.0	141	HBA_TRAFT	P01928 cebus apell
15	72	80.0	141	HBA_TRAFT	P01928 cebus apell
16	72	80.0	141	HBA_TRAFT	P01928 cebus apell
17	72	80.0	141	HBA_TRAFT	P01928 cebus apell
18	72	80.0	141	HBA_TRAFT	P01928 cebus apell
19	72	80.0	141	HBA_TRAFT	P18973 crocuta cro
20	72	80.0	141	HBA_TRAFT	P01959 equus asinu
21	72	80.0	141	HBA_TRAFT	P01961 equus hemio
22	72	80.0	141	HBA_TRAFT	P01960 equus asinu
23	72	80.0	141	HBA_TRAFT	P01960 equus asinu
24	72	80.0	141	HBA_TRAFT	P01960 equus asinu
25	72	80.0	141	HBA_TRAFT	P01960 equus asinu
26	72	80.0	141	HBA_TRAFT	P01960 equus asinu
27	72	80.0	141	HBA_TRAFT	P01960 equus asinu
28	72	80.0	141	HBA_TRAFT	P01960 equus asinu
29	72	80.0	141	HBA_TRAFT	P01960 equus asinu
30	72	80.0	141	HBA_TRAFT	P01960 equus asinu
31	72	80.0	141	HBA_TRAFT	P01960 equus asinu
32	72	80.0	141	HBA_TRAFT	P01960 equus asinu
33	72	80.0	141	HBA_TRAFT	P01960 equus asinu

34	72	80.0	141	HBA_MACCA	P21766 macaca assa
35	72	80.0	141	HBA_MACCA	P21767 macaca fasc
36	72	80.0	141	HBA_MACCA	P24659 macaca fasc
37	72	80.0	141	HBA_MACCA	P01925 macaca muli
38	72	80.0	141	HBA_MACCA	P19002 macaca mulo
39	72	80.0	141	HBA_MACCA	P21768 macaca mulo
40	72	80.0	141	HBA_MACCA	P07402 macaca mulo
41	72	80.0	141	HBA_MACCA	P11751 macaca mulo
42	72	80.0	141	HBA_MACCA	P01953 macaca mulo
43	72	80.0	141	HBA_MACCA	P23600 macaca mulo
44	72	80.0	141	HBA_MACCA	P20243 macaca mulo
45	72	80.0	141	HBA_MACCA	P23601 macaca mulo

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	140 AA
HBA_TRAFT	P04237			
AC	20-MAR-1987 (rel. 04, last sequence update)			
DT	20-MAR-1987 (rel. 04, last sequence update)			
PT	01-JUN-1994 (rel. 29, last annotation update)			
DE	HEMOGLOBIN ALPHA CHAIN.			
OS	Trachelaphus strepsiceros (Greater kudu).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Tragelaphus.			
OX	NCBI_TaxID=9446;			
RN	[1]			
FP	SEQUENCE.			
PX	MTG18E-95272623; P04237-4076993;			
RA	Edwards K., Wiesner H., Braunitzer G.;			
RT	"Primary structure of the hemoglobins from the greater kudu antelope			
PT	(Trachelaphus strepsiceros).";			
RL	Biol. Chem. Biophys. 266:395-402(1985).			
DR	PIR: A03293; UNPKA.			
DR	HSSP: P01966; UNPKA.			
DR	InterPro: IPR000971;			
DR	InterPro: IPR002318;			
DR	PIR: P00042; globin; 1.			
DE	PRINTS: PF00412; ALPHAHEM.			
DR	PROSITE: PS01033; GLOBIN; 1.			
KN	Heme; Oxygen transport; Respiratory protein; Erythrocyte.			
ET	METAL			
FT	57 57			
FT	86 86			
SO	SEQUENCE 140 AA; 14949 MW; AAD1FA85F4DE244 CRO64;			
Query Match	80.0%	Score 72;	DR 1;	Length 140;
Best local Similarity	100.0%	Pred. No. 4	86.06;	
Mismatches	13;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
QY	2 FPHFDLSHSAQVC 14			
DB	42 FPHFDLSHSAQVC 54			
RESULT	2			
HBA_TRAFT	STANDARD;	PRT:	141 AA.	
AC	P11755;			
DE	01-JUN-1994 (rel. 12, last sequence update)			
DE	01-JUN-1994 (rel. 29, last annotation update)			
OS	HEMOGLOBIN ALPHA-1 CHAIN.			
OS	Trachelaphus strepsiceros (Greater kudu).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Tragelaphus.			
OX	NCBI_TaxID=9446;			
RN	[1]			





```

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR HSP: A02295; HAEMK.
DR HSP: P01966; IHD.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00342; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte-
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15107 MW; 85AD233939EPFAD CRR64;

Query Match 80.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 9 9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 2 FPHFUSHGSAQV 14
DB 43 FPHFUSHGSAQV 55

RESULT 6
HBA_ANTPA STANDARD: PRT: 141 AA.
AC P1487;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Antrozous pallidus (pallid bat).
OC Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Placentalia;
OC Mammalia; Eutheria; Chordata; Mammalia; Eutheria; Placentalia;
OC Antrozous.
OX NCBI_TaxId=9440;
RN 11
RP SEQUENCE.
KW MEDLINE=86050104; PubMed=3675871;
KA Kleinschmidt T, Koop R F, Brannstrom G;
RT "The primary structure of the pallid bat (Antrozous pallidus,
RT Chiroptera) hemoglobin."
KL Biol. Chem. Hoppe-Seyler 368:1197-1202(1987).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR HSP: A02292; A29702.
DR HSP: P01922; IHD.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte-
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15107 MW; F66256668497646 CRR64;

```

```

HBA_ATEGE STANDARD: PRT: 141 AA.
ID HBA_ATEGE
AC P01927;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Ateles geoffroyi (Black-headed spider monkey), and
OS Saquinus oedipus (Cotton top tamarin).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates, Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxId=9509; 9490;
RN 11
RP SEQUENCE.
KW MEDLINE=75059126; PubMed=4215720;
FA Matsuda G, Maita T, Suenaga Y, Setoyuchi M, Ota Y, Araya A,
FA Goodman M, Barnabas J, Prychodko W;
ET "Studies on the primary structures of alpha and beta polypeptide
ET chains of adult hemoglobin of the spider monkey (Ateles geoffroyi).";
KL Hoppe Seyler's Z. Physik. Chem. 354:1517-1520(1978)
RN 121
RP SEQUENCE.
KW SPECIES=S.oedipus;
FX MEDLINE=8412383; PubMed=6437302;
FA Maita T, Hayashida M, Matsuda G;
RT "Primary structures of adult hemoglobins of silver marmoset,
RT Callithrix jacchus, and cotton-headed tamarin, Saquinus oedipus."
KL J. Biochem. 95:805-813(1984).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR HSP: A02253; HAEMK.
DR HSP: B28865; B28865.
DR HSP: P01922; IHD.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte-
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15039 MW; 387025A3DE88702 CRR64;

Query Match 80.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 9 9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 2 FPHFUSHGSAQV 14
DB 43 FPHFUSHGSAQV 55

RESULT 8
HBA_BISBO STANDARD: PRT: 141 AA.
ID HBA_BISBO
AC P09423;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA-1 AND ALPHA-11 CHAINS.
OS Bison boudaus (European bison).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovidae;
OX NCBI_TaxId=9502;
RN 11
RP SEQUENCE.
KW MEDLINE=86296173; PubMed=3741621;

```



```

DI 01-NOV-1990 (Rel. 16, Created)
DI 01-NOV-1990 (Rel. 16, Last sequence update)
DI 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Gallitrix argentea (Black-tailed marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrix.
OX NCBI_TaxId=9482;
RN [1]
RP SEQUENCE.
RX MEDLINE=8421384; PubMed=647202;
RA Matsuda T., Hayashi M., Matsuda G.;
RT "Primary structures of adult hemoglobins of silvery marmoset,
RT Callitrix jacchus, and cotton-headed tamarin, Saguinus oedipus."
RT J. Biochem. 95:805-813(1984).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: S06512; HSCUR.
DR HSP: P01922; IPDH.
DR InterPro: IPR000971; -.
DR InterPro: IPR002348; -.
DR Pfam: PF00042; globin.1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme oxygen transport, respiratory protein, erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15138 MW; 26B1307E2A68602 CRC64;

Query Match 80.0% Score 72; E: 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 9 aa-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 2 FPHPLSHGSAOV 14
DB 43 FPHPLSHGSAOV 55

RESULT 12
HBA_CAMDR STANDARD; PRI: 141 AA.
AC P01974;
DI 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Camelus dromedarius (dromedary) (Arabian camel), and
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Artiodactyla; Tylopoda; Camelidae; Camelus
OC NCBI_TaxId=9838; 9837;
RN [1]
RP SEQUENCE.
RX MEDLINE=80114125; PubMed=527943;
RA Braunlitzer G., Schrank B., Stangl A., Wiesner H.;
RT "Respiration at high altitudes, phosphate-protein interaction: the
RT sequence of hemoglobins from quinea pig and dromedary."
RT J. High-Altitude Biol. Physiol. Chem. 360:1941-1946(1979).
RN [2]
RP SEQUENCE.
RX SPECIES=C.dromedarius;
RA Braunlitzer G., Schrank B., Stangl A., Wiesner H.;
RT "Respiration at high altitudes, phosphate-protein interaction:
RT the sequence of the hemoglobins of the hamster (Mesocricetus auratus) and
RT the camel (Camelus ferus, Camelidae).";
RL J. Chem. Soc. Pak. 2:1-7(1980).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE

```

```

CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: B92758; HSCMA.
DR PIR: B91685; HSCMA.
DR HSP: P01955; 2ECH.
DR InterPro: IPR000971; -.
DR InterPro: IPR002348; -.
DR Pfam: PF00042; globin.1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme oxygen transport, respiratory protein, erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15172 MW; 269D80CF68676E5C CRC64;

Query Match 80.0% Score 72; E: 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 9 aa-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 2 FPHPLSHGSAOV 14
DB 43 FPHPLSHGSAOV 55

RESULT 13
HBA_CAMDR STANDARD; PRI: 141 AA.
AC P01970;
DI 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 30-MAY-2000 (Rel. 39, Last annotation update)
DE HEMOGLOBIN ALPHA 1 AND ALPHA 2 CHAINS.
OS Capra hircus (goat), Ammotragus lervia (Barbary sheep) (Mondak), and
OS Capra aries (sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxId=9925; 9999; 9940;
RN [1]
RP SEQUENCE FROM N.A. (ALPHA-1 AND -2).
RX SPECIES=C.hircus;
RX MEDLINE=82214818; PubMed=6282825;
RA Selton E.A., Wetliuk S.M., Linquell J.B.;
RT "Gene conversion of two functional goat alpha-globin genes preserves
RT only minimal flanking sequences."
RT J. Biol. Chem. 257:6845-6845(1982).
RN [2]
RP PARTIAL SEQUENCE.
RX SPECIES=C.hircus;
RX MEDLINE=68313123; PubMed=5658544;
RA Huisman T.H.J., Brandt G., Wilson J.B.;
RT "The structure of goat hemoglobins. II. Structural studies of the
RT alpha chains of the hemoglobins A and B."
RL J. Biol. Chem. 243:3675-3686(1968).
RN [3]
RP SEQUENCE.
RX SPECIES=A.lervia;
RX MEDLINE=7016388; PubMed=5449729;
RA Wilson J.B., Wrightstone R.N., Huisman T.H.J.;
RT "Hemoglobin alpha chain duplication in Barbary sheep, Ammotragus
RT lervia, Pallas, 1777."
RL Nature 226:354-355(1970).
RN [4]
RP SEQUENCE.
RX SPECIES=Ovis montanus;
RA Wilson J.B., Wrightstone R.N., Huisman T.H.J.;
RT "A partial amino acid sequence for sheep haemoglobin A."
RL Biochem. J. 103:129-140(1967).
RN [5]

```



Query Match: 83.0%; Score 72; DB L: Length 141;  
 Best Local Similarity: 100.0%; Pval No: 9 9e-06;  
 Matches: 33; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;  
 QY: 2 PPHFISHGSAOV 14  
 |||||  
 DB: 43 PPHFISHGSAOV 55

Search completed: June 21, 2001, 10:29:20  
 Job time: 249 sec

.

.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compaq Inc.

OM protein - protein search, using SW model

Run on: June 21, 2001, 10:29:04 ; Search time 33.64 seconds

(without alignments)  
58,995 Million cell updates/sec

Title: US-08-832-443b-2

Perfect score: 90

Sequence: 1 GPHFDLSHSAQVC 15

Scoring table: GPHFDLSHSAQVC 15

Searched: 425026 seqs, 13305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: SP\_TREMBL\_16: \*  
2: SP\_Archea: \*  
3: SP\_Bacteria: \*  
4: SP\_Fungi: \*  
5: SP\_Human: \*  
6: SP\_Invertebrate: \*  
7: SP\_Misc: \*  
8: SP\_Mammal: \*  
9: SP\_Organelle: \*  
10: SP\_Phage: \*  
11: SP\_Plant: \*  
12: SP\_Rodent: \*  
13: SP\_Unclassified: \*  
14: SP\_Virus: \*

Pred. No. is the number of results predicted by chance. To have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	80.0	141	6	Q9TS35
2	72	80.0	141	6	Q9TS34
3	72	80.0	142	4	Q9NVR7
4	72	80.0	142	4	Q28744
5	72	80.0	142	6	Q28744
6	72	80.0	142	6	Q28745
7	72	80.0	142	6	Q28383
8	72	80.0	142	6	Q9XSN3
9	72	80.0	142	6	Q9XSN2
10	72	80.0	142	6	Q9XSK1
11	72	80.0	142	6	Q9XSE9
12	72	80.0	142	6	Q9XSE8
13	72	80.0	142	6	Q9TVA4
14	72	80.0	142	6	Q9TVA3
15	72	80.0	142	6	Q9TTS1
16	72	80.0	142	6	Q9TTS0
17	72	80.0	142	6	Q9TTR9
18	72	80.0	142	6	Q9TSN9
19	72	80.0	142	6	Q9TSN8

20	72	80.0	142	6	Q9TSN7	Q9TSN7 bubalus bub
21	72	80.0	142	6	Q9TNR6	Q9TNR6 equus asinu
22	71	78.9	100	13	Q9W068	Q9W068 phasianus c
23	71	78.9	100	13	Q9W073	Q9W073 gallus gall
24	71	78.9	136	13	Q91249	Q91249 columba sp.
25	71	79.9	142	13	Q12986	Q12986 columba liv
26	69	76.7	58	11	Q61649	Q61649 mus musculu
27	69	76.7	122	11	Q90803	Q90803 mus musculu
28	69	76.7	142	11	Q61287	Q61287 mus musculu
29	61	67.8	140	13	Q12985	Q12985 columba liv
30	54	60.0	141	13	Q90825	Q90825 aequidone
31	54	60.0	142	13	Q91Y02	Q91Y02 byctidius re
32	53	58.9	99	11	Q61654	Q61654 mus musculu
33	50	55.6	458	4	Q9NVR4	Q9NVR4 homo sapien
34	50	55.6	648	4	Q9P209	Q9P209 homo sapien
35	49	54.4	141	13	Q9P819	Q9P819 naja naja (
36	48	53.3	142	11	Q63910	Q63910 rattus norv
37	44	48.9	556	10	Q9M172	Q9M172 aradidopsis
38	43	47.8	172	5	Q25347	Q25347 leishmania
39	43	47.8	635	4	Q13476	Q13476 homo sapien
40	42	46.7	536	5	Q91Y02	Q91Y02 caenorhabdi
41	42	46.7	838	5	P90882	P90882 caenorhabdi
42	42	46.7	1101	10	Q92907	Q92907 cyrtus sativ
43	42	46.7	2148	11	Q93505	Q93505 rattus norv
44	41.5	46.1	581	10	Q9SX39	Q9SX39 aradidopsis
45	41.5	46.1	1376	4	P82279	P82279 homo sapien

## ALIGNMENTS

RESULT 1  
ID Q9TS35 PRELIMINARY: PRT: 141 AA.

AC Q9TS35;  
DT 01-MAY-2000 (TREMBLrel. 13, created)  
DI 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DI 01-MAR-2001 (TREMBLrel. 16, last annotation update)  
DE ALPHA 1 GLOBIN.  
OS Hylobates lar (Common gibbon).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates;  
OX NCBI\_TaxID:9560;  
RN [1]  
RP SEQUENCE FROM N.A.  
EX MEDLINE:92406743; PubMed:1526980;  
RA Bailey A.D., Stanhope M., Sligham J.L., Goodman M., Shen C.C.,  
FA Shen C.K.;  
RT "Randomly duplicated alpha globin genes of gibbon.";  
RL J. Biol. Chem. 267:18398-18406(1992).  
CC -1- SIMILARITY: TO GLOBIN FAMILY.  
CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.  
DE HESRP: P01922; 1820.  
DR InterPro: IPR000971;  
DR InterPro: IPR002338;  
DR Pfam: PF00642; globin: 1.  
DR PRINTS: PR00012; ALPHABEM.  
DR PROSITE: PS01933, GLOBIN, 1.  
KW Erythrocyte, Heme, Oxygen Transport, Respiratory protein,  
SEQUENCE 141 AA, 15136 MW, 2913618 Da, 7706 kDa, 7706 kDa

Query Match: 83.0%, Score 72, DB 6, Length 141;  
Best local Similarity: 100.0%, P-Val: 2.4e-05;  
Matches: 13; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 2 GPHFDLSHSAQVC 14  
DB 43 GPHFDLSHSAQVC 55

RESULT 2  
Q9TS34





DR HSSP: P01966; 1HBA.  
 DR Interpro: IPR000971; ..  
 DR Interpro: IPR002388; ..  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01033; GLOBIN\_1.  
 DR Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 KW Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 SO SEQUENCE 142 AA; 15236 MW; 07094DFFR87033 CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHDLSHSAOV 14  
 ID 028745  
 AC 028745  
 DB 44 FPHDLSHSAOV 56

RESULT 6  
 ID 028745 PRELIMINARY; PFT; 142 AA.  
 AC 028745;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE ALPHA GLOBIN CHAIN.  
 OS Ovis aries (Sheep).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Kunitaria; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Ovis.  
 NX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 FX MFTIME=9528794; PubMed 776611;  
 RA Rietveld M S., Casula S., Rando A., Vestri P.  
 RT "Sheep alpha-globin gene sequence: implications for their concerted  
 evolution and for the down-regulation of the 3' genes".  
 RL J. Mol. Evol. 40:449-453(1995).  
 CC -1- SIMILARITY: TO GLOBIN FAMILY.  
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.  
 DR EMBL: X79214; AAA4751.1; ..  
 DR HSSP: P01966; 1HBA.  
 DR Interpro: IPR000971; ..  
 DR Interpro: IPR002388; ..  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01033; GLOBIN\_1.  
 KW Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 SO SEQUENCE 142 AA; 15178 MW; 07094DFFR87033 CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHDLSHSAOV 14  
 ID 028745  
 AC 028745  
 DB 44 FPHDLSHSAOV 56

RESULT 7  
 ID 028745 PRELIMINARY; PFT; 142 AA.  
 AC 028745;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE B11 ALBA-2 CHAIN.  
 OS Equus caballus (Horse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 NX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 FX MFTIME=9528794; PubMed 776611;  
 RA Rietveld M S., Casula S., Rando A., Vestri P.  
 RT "Sheep alpha-globin gene sequence: implications for their concerted  
 evolution and for the down-regulation of the 3' genes".  
 RL J. Mol. Evol. 40:449-453(1995).  
 CC -1- SIMILARITY: TO GLOBIN FAMILY.  
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.  
 DR EMBL: X79214; AAA4751.1; ..  
 DR HSSP: P01966; 1HBA.  
 DR Interpro: IPR000971; ..  
 DR Interpro: IPR002388; ..  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01033; GLOBIN\_1.  
 KW Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 SO SEQUENCE 142 AA; 15220 MW; 979F83BD040317E0 CR664;

OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=98215145; PubMed 2815578;  
 FX Clegg J.B.;  
 RT "Gene conversions in the horse alpha globin gene complex".  
 PL Mol. Biol. Evol. 4:492-503(1987).  
 PN [2]  
 RP SEQUENCE FROM N.A.  
 RA Clegg J.B.;  
 RT "Schwartz (1971-1994) to the PubMed/Entrez/Refseq databases.  
 CC -1- SIMILARITY: TO GLOBIN FAMILY.  
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.  
 DR EMBL: M17901; AAA3947.1; ..  
 DR HSSP: P01959; 1HBE.  
 DR Interpro: IPR000971; ..  
 DR Interpro: IPR002388; ..  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01033; GLOBIN\_1.  
 KW Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 SO SEQUENCE 142 AA; 15229 MW; 93AP5H9B50B0B1 CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHDLSHSAOV 14  
 ID 028745  
 AC 028745  
 DB 44 FPHDLSHSAOV 56

RESULT 8  
 ID 028745 PRELIMINARY; PFT; 142 AA.  
 AC 028745;  
 DT 01-NOV-1996 (TrEMBLrel. 12, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE ALPHA 1 HEMOGLOBIN.  
 OS Equus caballus (Horse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 NX NCBI\_TaxID=9790;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99055770; PubMed 9847419;  
 FX Oakenfull E.A., Clegg J.B.;  
 RT "Phylogenetic relationships within the equus genus and the evolution  
 of alpha and beta globin genes".  
 RL J. Mol. Evol. 47:772-783(1998).  
 CC -1- SIMILARITY: TO GLOBIN FAMILY.  
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.  
 DR EMBL: U79200; AAB94461.1; ..  
 DR HSSP: P01958; 1HBE.  
 DR Interpro: IPR000971; ..  
 DR Interpro: IPR002388; ..  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01033; GLOBIN\_1.  
 KW Erythrocyte Heme Oxygenase Transferrin Respiratory Protein.  
 SO SEQUENCE 142 AA; 15220 MW; 979F83BD040317E0 CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPHDLSHSAOV 14  
 ID 028745  
 AC 028745  
 DB 44 FPHDLSHSAOV 56



PA MEDLINE:99065770; PubMed:9847619;  
 RA Oakenfull E.A., Clegh J.B.;  
 RT "Physico-chemical relationships within the genus *Equus* and the evolution  
 of alpha and theta globin genes.";  
 RL J. Mol. Evol. 47:772-783(1998).  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RA Oakenfull E.A., Clegh J.B.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO GLOBIN FAMILY.  
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.  
 DE EMBL: A4329439, AAB93466.1;  
 DE EMBL: A4329138, AAB93466.1; JOINED.  
 DR HSP: P01936; 118E.  
 DR Interpro: IPRO00971; 1;  
 DR Interpro: IPRO02338; 1;  
 DR Pfam: PF00042; Globin\_1.  
 DR PRINTS: PR00612; ALPHAHEM.  
 DR PROSITE: PS01633; GLOBIN\_1  
 KW Erythrocyte; Hemoglobin; Oxygen transport; Respiratory protein.  
 SO SQUONCH 142 AA; 1521 MW; 92kDa;P01934-9824 GENE41;

```

Query Match 72: 18 6: Length 142:
Best Local Similarity 100.0%
Matches 15, Conservative 0, Mismatches 0, Indels 0
OY 2 PFHFDLSSGSAOV 14
| | | | | | | | | |
DB 44 PFHFDLSSGSAOV 56

```

RESULT	13
Q9TV44	
ID	Q9TV44
AC	Q9TV44
DT	01-MAY-2006 (TFPMRel_13) Created
DT	01-MAY-2006 (TFPMRel_13) Last sequence update
DT	01-MAY-2006 (TFPMRel_16) Last annotation update
DE	ALPHA 1 HEMOGLOBIN.
OS	Equus hemionus (Kulan).
OC	Eukaryota; Metazoa; Chordata; Graziata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX	NCBI_TaxID=9794;
FN	[1]
RP	SEQUENCE FROM N.A.
FX	MEF1NF 9066770; P08MD-9947419;
RA	Oakenf, F.A., Glegg J.B.;
RT	"Phylogenetic relationships within the genus Equus and the evolution
RL	of alpha and theta globin genes.";
	J. Mol. Evol. 47:772-783(1998).

Query Match	80.0%;	Score 72;	DB 6;	Length 142;
Best Local Similarity	100.0%;	Pred. No. 2.4e-05;		
Matches	13;	Conserved+120	0;	Mismatches+400
			0;	Indels
			0;	Gaps

O <sub>2</sub>	2	PPHPII.SHGSAQV	14
Ph	44	PPHPII.SHGSAQV	56

RESULT	14			
O9TVA3				
10	O9TVA3	PRELIMINARY:	PR1:	142 AA.
AC	O9TVA3			
DT	01-MAY-2000 (ITEMBLREL_13, Created)			
DT	01-MAY-2000 (ITEMBLREL_13, Last sequence update)			
DT	01-MAY-2001 (ITEMBLREL_15, Last annotation update)			
DE	ALPHA 2 HEMOGLOBIN.			
OS	Equus burchelli (Plains zebra) (Equus quagga).			
OS	Equus burchelli (Plains zebra) (Equus quagga).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX	NCBI-TAXID=9790;			
RN	11			
RP	SEQUENCE FROM N.A.			
FX	MEDLINE 99065770; PubMed 9947419;			
RA	Cabotelli E.A., Claga J.B.;			
RT	"Phylogenetic relationships within the genus Equus and the evolution			
KL	of alpha and theta globin genes.";			
RN	J. Mol. Evol. 47:772-783(1998).			
RP	12			
RP	SEQUENCE FROM N.A.			
RA	Cabotelli E.A., Claga J.B.;			
PL	Submitted (FBI-1999) to the EMBL/Genbank/DBJ databases.			
CC	-1- SIMILARITY: TO GLOBIN FAMILY.			
CC	-1- SIMILARITY: TO ALPHA HAHM-GLOBIN FAMILY.			
DR	EMBL: AF129135; AAB93462.1;			
DR	EMBL: AF129134; AAB93462.1; JOINED.			
DR	HSSP: P01958; 11RF.			
DR	InterPro: IPR000971;			
DR	InterPro: IPR002388;			
DR	PIRMS: PFO0042; GLOBIN_1.			
DR	PRINTS: PR00612; ALPHAHAM.			
DR	PROSITE: PS01033; GLOBIN_1.			
DR	Protein-tyrosine-kinase oxygen transport, respiratory protein.			
SO	SEQUENCE: 142 AA; 15190 MW; 627837RC00010Bb5; Trc94;			

Query Match	(8), (9); Score: 72; NR 6; Evalue: 142;
Post Local Similarity	100.0%; Pred. NO. 2.4e-05;
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT	15	PEPCLIMINABV:	PRT:	142 AA.
09TST1				
10	Q9TST1			
AC	Q9TST1			
DT	01-MAY-2000 (TTRMBREL_13, created)			
DT	01-MAY-2000 (TTRMBREL_13, last sequence update)			
DT	01-MAY-2001 (TTRMBREL_16, last annotation update)			
TE	HEMOGLOBIN ALPHA CHAIN.			
CN	ALPHA GLOBIN.			
OS	Bos taurus (Hovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Moleculostomi;			
OC	Mammalia; Eumetazoa; Cetartiodactyla; Eumammalia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PODOLIAN.			
RA	Fuller R., Phlorastoin F., Vinrenti D., Campanile C., Di Iorio A.;			
RT	"Nucleotide sequence of Podolian cattle (Bos taurus primigenius) alpha			
RT	globin genes."			
RI	Submitted (MAY-1999) to the EMBL/Genbank/Trna databases.			
CC	- SIMILARITY: TO GLOBIN FAMILY			

RT	"Nucleotide sequence of Fodolian cattle ( <i>Bos taurus primigenius</i> ) alpha globin genes."
RT	Globin genes.
FI	Cloned (MAY 1999) as the 1997/2000 data bases.
CC	1- SIMILARITY: TO GLOBIN FAMILY.

